

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: March 2, 2006, 06:32:30 ; Search time 1152 Seconds
(without alignments) 799.304 Million cell updates/sec

Title: US-10-041-030-4

Perfect score: 2290
Sequence: 1 MRSPPGEHCAPKPEPVXYG.....ATQLVGEQNCIKLIPGSPID 420

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7218512 segs, 1096188070 residues

Total number of hits satisfying chosen parameters: 14437024

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlh
-Q=/abs/ABSSMB.spool/US10041030/runat_01032006_134413_22030/app_query.fasta.1
-DB=Published Applications NA New -QMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=humand0.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs805h
-USER=US10041030 @CEN 1.1 220 @runat 01032006 134413 22030 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOC -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRBADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA New:

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1482.5	64.7	2508	9	US-11-072-512-1958 Sequence 1958, Ap
2	711.5	31.1	1105	8	US-10-750-185-40960 Sequence 40960, A
3	711.5	31.1	1105	8	US-10-750-623-40960 Sequence 40960, A
4	293	12.8	1345	6	US-09-925-065A-679103 Sequence 679103,

5	102	4.5	3078	8	US-10-750-185-38950	Sequence 38950, A
6	102	4.5 <td>3078</td> <td>8<td>US-10-750-623-38950</td><td>Sequence 38950, A</td></td>	3078	8 <td>US-10-750-623-38950</td> <td>Sequence 38950, A</td>	US-10-750-623-38950	Sequence 38950, A
7	101.5	4.4	119160	12 <td>US-11-121-086-12</td> <td>Sequence 12, Appl</td>	US-11-121-086-12	Sequence 12, Appl
8	101.5	4.4	167891	12 <td>US-11-121-086-14</td> <td>Sequence 14, Appl</td>	US-11-121-086-14	Sequence 14, Appl
9	100.5	4.4	1664	12 <td>US-11-013-247A-16</td> <td>Sequence 16, Appl</td>	US-11-013-247A-16	Sequence 16, Appl
10	97.5	4.3	164810	12 <td>US-11-121-086-4</td> <td>Sequence 4, Appl</td>	US-11-121-086-4	Sequence 4, Appl
11	97	4.2	21728	7 <td>US-10-330-773-362</td> <td>Sequence 362, Appl</td>	US-10-330-773-362	Sequence 362, Appl
12	97	4.2	150468	12 <td>US-11-112-908-56</td> <td>Sequence 56, Appl</td>	US-11-112-908-56	Sequence 56, Appl
13	97	4.2	193789	12 <td>US-11-112-908-55</td> <td>Sequence 55, Appl</td>	US-11-112-908-55	Sequence 55, Appl
14	96.5	4.2	1770	12 <td>US-11-183-136-41</td> <td>Sequence 41, Appl</td>	US-11-183-136-41	Sequence 41, Appl
15	94.5	4.1	1363	8 <td>US-10-750-185-32846</td> <td>Sequence 32846, A</td>	US-10-750-185-32846	Sequence 32846, A
16	94.5	4.1	1363	8 <td>US-10-750-623-32846</td> <td>Sequence 32846, A</td>	US-10-750-623-32846	Sequence 32846, A
17	94.5	4.1	3510	9 <td>US-11-114-962-9</td> <td>Sequence 9, Appl</td>	US-11-114-962-9	Sequence 9, Appl
18	94	4.1	3430	8 <td>US-10-750-185-62790</td> <td>Sequence 62790, A</td>	US-10-750-185-62790	Sequence 62790, A
19	94	4.1	3430	8 <td>US-10-750-623-62790</td> <td>Sequence 62790, A</td>	US-10-750-623-62790	Sequence 62790, A
20	94	4.1	53916	7 <td>US-10-330-773-65</td> <td>Sequence 65, Appl</td>	US-10-330-773-65	Sequence 65, Appl
21	93.5	4.1	1400	12 <td>US-11-136-527-5713</td> <td>Sequence 5713, Ap</td>	US-11-136-527-5713	Sequence 5713, Ap
22	93.5	4.1	1650	12 <td>US-11-013-247A-1</td> <td>Sequence 1, Appl</td>	US-11-013-247A-1	Sequence 1, Appl
23	93.5	4.1	2475	12 <td>US-11-136-527-1617</td> <td>Sequence 1617, Ap</td>	US-11-136-527-1617	Sequence 1617, Ap
24	93.5	4.1	6510	8 <td>US-10-995-561-112</td> <td>Sequence 112, Appl</td>	US-10-995-561-112	Sequence 112, Appl
25	93.5	4.1	7823	8 <td>US-10-995-561-117</td> <td>Sequence 117, Appl</td>	US-10-995-561-117	Sequence 117, Appl
26	93	4.1	1224	8 <td>US-10-750-185-45251</td> <td>Sequence 45251, A</td>	US-10-750-185-45251	Sequence 45251, A
27	93	4.1	1224	8 <td>US-09-925-065A-458917</td> <td>Sequence 458917, A</td>	US-09-925-065A-458917	Sequence 458917, A
28	92.5	4.0	590	6 <td>US-10-787-906-12</td> <td>Sequence 12, Appl</td>	US-10-787-906-12	Sequence 12, Appl
29	92	4.0	1560	7 <td>US-11-072-512-1600</td> <td>Sequence 1600, Ap</td>	US-11-072-512-1600	Sequence 1600, Ap
30	92	4.0	2483	9 <td>US-10-821-234-43</td> <td>Sequence 43, Appl</td>	US-10-821-234-43	Sequence 43, Appl
31	92	4.0	2767	8 <td>US-11-136-527-2243</td> <td>Sequence 2243, Ap</td>	US-11-136-527-2243	Sequence 2243, Ap
32	92	4.0	2936	12 <td>US-11-136-527-2243</td> <td>Sequence 25, Appl</td>	US-11-136-527-2243	Sequence 25, Appl
33	92	4.0	5928	8 <td>US-10-863-093-25</td> <td>Sequence 29, Appl</td>	US-10-863-093-25	Sequence 29, Appl
34	91.5	4.0	1569	12 <td>US-11-076-427A-29</td> <td>Sequence 27, Appl</td>	US-11-076-427A-29	Sequence 27, Appl
35	91.5	4.0	1571	12 <td>US-11-076-427A-27</td> <td>Sequence 3, Appl</td>	US-11-076-427A-27	Sequence 3, Appl
36	91.5	4.0	7893	12 <td>US-11-186-731-3</td> <td>Sequence 1, Appl</td>	US-11-186-731-3	Sequence 1, Appl
37	91.5	4.0	8106	12 <td>US-11-186-731-1</td> <td>Sequence 769, Appl</td>	US-11-186-731-1	Sequence 769, Appl
38	91.5	4.0	8250	8 <td>US-10-453-372-769</td> <td>Sequence 6, Appl</td>	US-10-453-372-769	Sequence 6, Appl
39	91.5	4.0	23907	12 <td>US-11-186-731-6</td> <td>Sequence 40, Appl</td>	US-11-186-731-6	Sequence 40, Appl
40	91.5	4.0	24230	12 <td>US-11-186-731-4</td> <td>Sequence 2936, Ap</td>	US-11-186-731-4	Sequence 2936, Ap
41	91.5	4.0	162289	12 <td>US-11-121-086-20</td> <td>Sequence 258, Appl</td>	US-11-121-086-20	Sequence 258, Appl
42	91	4.0	2156	12 <td>US-11-136-527-2936</td> <td>Sequence 551509, A</td>	US-11-136-527-2936	Sequence 551509, A
43	90.5	4.0	2736	9 <td>US-11-072-512-258</td> <td>Sequence 551510, A</td>	US-11-072-512-258	Sequence 551510, A
44	90	3.9	1970	6 <td>US-09-925-065A-551509</td> <td></td>	US-09-925-065A-551509	
45	90	3.9	1970	6 <td>US-09-925-065A-551510</td> <td></td>	US-09-925-065A-551510	

RESULT 1
US-11-072-512-1958
Publication No. US20060029945A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: HIO, YUUKO
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHITO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOMYUKI
APPLICANT: NAGAHARA, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072, 512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350, 978

	PRIOR FILING DATE:	2002-01-25
	PRIOR APPLICATION NUMBER:	JP 2001-379298
	PRIOR FILING DATE:	2001-11-05
	NUMBER OF SEQ ID NOS:	4096
	SOFTWARE:	Patentin Ver. 2.1
	SEQ ID NO	1958
	LENGTH:	2508
	TYPE:	DNA
	ORGANISM:	Homo sapiens
	US-11-072-512-1958	
	Alignment Scores:	
	Pred. NO.:	5.33e-127
	Score:	1482.50
	Percent Similarity:	74.2%
	Best Local Similarity:	65.6%
	Query Match:	64.7%
	DB:	9 Gaps: 4
US-10-041-030-4 (1-420) x US-11-072-512-1958 (1-2508)		
QY	3 SerProglIyGLInGluGlnHIScySaLaPrOaeMlyGluProValIySTyGIyGluLeu	22
Db	208 TCTCCCGGT-----GAGATGGCAGCCAGCGCAAGAGGCCATTCACAGTATGGTAATC	261
QY	23 ValValLeuGLyTyraSnGLyAlaLeuProAenGLyAspArgLyLyrArgLySerArg	42
Db	261 -----	261
QY	43 PheAlaLeuTyrlYArGrProLySaLaenGLyValIyPSeSerThrValHISvalIle	62
Db	262 -----ATC	264
QY	63 SerThrProGLInALserIySaLaIeserCySlyGLyGlnHISerIleserTyThr	82
Db	265 TCACAGCGCGTGCTCCCAAGGCACAGATTAAcCGGTGCACAGACAGCATCTGTATACA	324
QY	83 LeuSerArgSaenGLnThrValValIaGLyTyThrHISapRlySaPrThraPMePhe	102
Db	325 CTGTCCCGAGGCACCTCGGTCTATAGTGAATACATGATAGCGACACAGCATGTTTC	384
QY	103 GLnValGLyArgSerThrGLuSerProIIeaSPheValValHISApThrIleserGLy	122
Db	385 CAGATTGGCGCGTCCACAGAGAACAGATTGACTTGTTGGTTAACAGACAGTCCCGTGA	444
QY	123 SerGLIneThraSpGLuaGLnIleThrGLInSerThrIleserAPheAlaCySalg	142
Db	445 GGA---GGGCGTCGCGAGCGCCCTTCTGCCACAGACACATCTCCGCTATGGCTGCCG	501
QY	143 ILeValCYsaSPArXaBnGLUProTyThrTrlaARGLIlePheAlaIaAGIYPheaSPer	162
Db	502 ATCTCTGTGACCGCGCGCACCTTACTGCCCCGATCTATGTCGGTGGCTTCGATGCC	561
QY	163 SerTyASnIlePheLeuGLyGLyValaIaLYeTyrlYsaBnPProABrgLyHISmet	182
Db	562 TCTACCAACATCTTCTCTTGGAGAGCGAGCGCCAATGGGGACCCCAAGTAGCGCTGATG	621
QY	183 AspgLyLeuThraSnGLyValIeUvaLMetHISProARGLyGLyPHeThrGLInGLu	202
Db	622 GATGACATGCACCACCATGGAATGCTGTGTGATGCACCCGCGACGCGCTTCTCCAGGAC	681
QY	203 SerGLInProGLIValITraprArgLIIseserValCYseGLYasPryalTYThrLeuARGLu	222
Db	682 TCAGCGCCGGGTGTCTGGCGGGAGATCTGGTCTGTGGAGATGTTHACATTTGGGGAC	741
QY	223 ThraYSerAIaGLInGLnaRGLyLyLeuValaGLuSerGLuThraSnValLeuGLInASP	242
Db	742 AGCGCTCAGCCACAGACGCGGGCAAGCTGTGTGAAAACGAGTCCAACGTGTGCAGGAC	801
QY	243 GlySerLeuIIleaPLeuCYseGLyAlaThrIleUeuThraPgrThralaaBrgLyLeuPhe	262
Db	802 GGCTCTTCATGACCTGTGTGTGGGGCAACATGCTGTGGGACACCGCGGGGCTGCTG	861

OY		263	HistHrProThrGlnIySNIeGlLualaleuArgInguLiLeaMAlALARPro	282
Db		862	CGGGCTCCCAACtGMAAGAACTGTGAAGGCCAAGGAGGAGGCAATAGCACGGCGCC	921
OY		283	GlnCyPProValGIyleuAsnThrlleuAlaBepProSerIlleaMaRgLySGluValAl	302
Db		922	CAGTGCCTCCGTTAGGAGCTTAGACACTTGACTTCCTCCCAAGCCAGCCCGTGGCCGACAGCG	981
OY		303	GluGluIyUv---GlnProTPPLAtLYtleusercYsgLIyNIsvalHISglTYTHIASN	321
Db		982	CCCCAACAAACGACAGCCCTGGAGTCTACGCTCGCGGSCACGTCCAGAGCTTACCACAGC	1041
OY		322	TrrpGLIHASgsArAPTrHrGluAlaBSngLIuRgGluCySPromerCYAARGThrVal	341
Db		1042	TGGGCTGCCCGGGAGCGGGGGCCCCCAGAGAGCGCAATATCTCTGTGGCGCTTGTG	1101
OY		342	GIlyProTYrValProLeuTrpLeugLIyCSgLIuAlagIyPhetyrValasPalagIyPro	361
Db		1102	GGGCTTAATRGCTCTATGTGCTTGCCAGAGAGCGCGCTCTGCGCTGAGACCTGGGCGG	1161
OY		362	ProThrHisAlaphethrProCyegLIyNIsvalCyssergLIySSeralILysTYrTrp	381
Db		1162	CCTAGCCATGCTTTGCACTTGACCTTGCGGCCACCTGTGCTGTAGAAGACATGCCCGCTACCGG	1221
OY		382	SerGlnIleProleuProHISglYTThrlHisAlaHeHISAlaAIaCYsProPhCYela	401
Db		1222	CGCCAGACACCATGCCCCACGACGCCATCCATCTTCATGCGCGCTGCCCCCTTTTGGCGGG	1281
OY		402	ThrgInleuValIGlYGluGlnAsnCysIIelySeullepehngLIyProIleASP	420
Db		1282	GCTGCGCTTACCGGCGAGCATGCGTGGCTCGGCTCANTTTCCAGAGGCGCGCTGAT	1338
RESULT 2				
US-10-750-185-40960/c				
; Sequence 40960, Application US/10750185				
; Publication No. US20050260603A1				
GENERAL INFORMATION:				
; APPLICANT: MMT GENOMICS, INC.				
; APPLICANT: DENISE, Sue K.				
; APPLICANT: KERR, Richard				
; APPLICANT: ROSENFELD, David				
; APPLICANT: HOLM, Tom				
; APPLICANT: BATES, Stephen				
; APPLICANT: PANTIN, Dennis				
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS				
; FILE REFERENCE: MM1100-2				
; CURRENT APPLICATION NUMBER: US/10/750,185				
; PRIOR FILING DATE: 2003-12-31				
; PRIOR APPLICATION NUMBER: US 60/437,482				
; NUMBER OF SEQ ID NOS: 64922				
; SOFTWARE: PatentIn version 3.1				
; SEQ ID NO 40960				
; LENGTH: 1105				
; TYPE: DNA				
; ORGANISM: Bovine				
US-10-750-185-40960				
Alignment Scores:				
Pred. No.:	4.41e-56	Length:	1105	
Score:	711.50	Matches:	164	
Percent Similarity:	66.9%	Conservative:	48	
Best Local Similarity:	51.7%	Mismatches:	97	
Query Match:	31.1%	Indels:	12	
DB:	8	Gaps:	4	
US-10-041-030-4 (1-420) x US-10-750-185-40960 (1-1105)				
OY		108	ThrgIuseProIleAspPheValThraSPThrlIeserGIySerGlnAsnThraSP	127
Db		1102	ACTGAAGTCATTAAATTTGTATAACTATATAGTTCAGAGAGT-AAAGATTTGAT	1044
OY		128	-----GluAlaGlnIleThrgInserThrlIeserArgPheAlaCYeArgIleValaCYs	145


```
QY 365 lapherProCyGlyVhIeValCySeRgIuLySeRAlaLyTyTrpSer-GlnIle 384
Db 335 TAGTAGACACCTGTGTAGCGAAGTTCAAGAAAAAATATCTCCATGAGCCCAATC 276
QY 385 ProLeuProHISgIuThrHISaIaRheHISaIaCyAProPheCyValaThgInLeu 404
Db 275 CCACTTCTCATGGTCTCATGTGTATGACAGCTGGCCCTTTGTGTCCATGAGTTG 216
QY 405 ValGIGluGlnIaMnCySleIyLeuIlePheGlnIlyProIle 419
Db 215 GTTGAGAACAGAGCTACATTAGCATTTAGTCAAGACCTTTA 171

RESULT 4
US-09-925-065A-679103
/ Sequence 679103, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925, 065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/269,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 679103
/ LENGTH: 1345
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-679103

Alignment Scores:
Pred. No.: 2,86e-17 Length: 1345
Score: 293.00 Matches: 80
Percent Similarity: 49.2% Conservative: 17
Best Local Similarity: 40.6% Mismatches: 50
Query Match: 12.8% Indels: 53
Gaps: 6
DB: 6

US-10-041-030-4 (1-420) x US-09-925-065A-679103 (1-1345)
QY 170 GlnLysAlaAlaIySTpLySaenProAaPGLYHISeAaPGLyLeuThrHISaNgly 189
Db 354 GAGCGAGCGGCGCAATAGCGAGCCCGAGATGCGCTGATGATGAGACACCAATGGA 413
QY 190 ValLeuValMetHisPAAaGgIyGlyPheThrGluGlnIuSerGlnProGlyValTTPArg 209
Db 414 GTCTGTGTATGACCCGCGAGCGGCTCTCCAGAGACTCAGCCCGGTGTCTGTGCGG 473
QY 210 GluIleSerValCySgIyAaPValTyTrHISaRgGlnIuThrAaSerAlaGlnIaXg 229
Db 474 GAATCTCGGTCGTGTGGAATGTGTACATTGCGGAGACAGCCCTCAGCCAGCAGCGG 533
QY 230 GILyLeuValGluSerGluThrAaenValLeuGlnAaPGLySerLeuIleAaPLeuCyS 249
Db 534 GCGAAGCTGGTA----- 545
QY 250 GILyAlaThrLeuLeuTTPArgThrAlaAaPGLyLeuPheHISgThrProThrGlnIyHIS 269
Db 546 -----GATGCGCCGCTCCATTC-CCACCC----- 568
* QY 270 IlegIaLeuAaRgGlnIuIleAaenAlaAlaAaRgProGlnCySProValGly----- 287
```

```
Db 569 ---CCTATCTTCCAGAGGCTTCAGAAAGTGGCCCAATCCCC-TCGCCAAGCCAGAGATG 624
QY 288 LeuAaThrLeuAlaPheProSerIleAaAaGlyGluValValGluIyGlnPro 307
Db 625 ATCTCCAGTCTTCTGTGCGGAGAGCCCAAGAGCTCCATTCAGAGAAATCCAGGCC 684
QY 308 TrpAlaTyIeUeSerCySgIyHISaValHISgIyTyHISaenIlePGLyHISaTsSerAaP 327
Db 685 CAGGGGCTTTTGCATGACCCACTTG-----TCACAGACAGAGC- 725
QY 328 ThrGluAlaAaNgIuAaRgIuCySProMetCyAaRgThrValGlyPro----- 343
Db 726 AGACAGACAGAGGTCCCAAGACATGACATGACATGACATGACAGAGGTCCAGAGGTGATGCG 785
QY 344 -----TyValProLeuITrLeu---GlyCyS 351
Db 786 ATCAAGGACACACAGACATCTGTGCAATCCAGAGTGGCTTCAGAGCTGC 836

RESULT 5
US-10-750-185-38950
/ Sequence 38950, Application US/10750185
/ Publication No. US20050260603A1
/ GENERAL INFORMATION:
/ APPLICANT: MMI GENOMICS, INC.
/ APPLICANT: DENISE, Sue K.
/ APPLICANT: KERR, Richard
/ APPLICANT: ROSENFIELD, David
/ APPLICANT: HOLM, Tom
/ APPLICANT: BATES, Stephen
/ APPLICANT: FANTIN, Dennis
/ TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
/ FILE REFERENCE: MM1100-2
/ CURRENT APPLICATION NUMBER: US/10/750,185
/ PRIOR FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437,482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 38950
/ LENGTH: 3078
/ TYPE: DNA
/ ORGANISM: Bos taurus
US-10-750-185-38950

Alignment Scores:
Pred. No.: 34.4 Length: 3078
Score: 102.00 Matches: 88
Percent Similarity: 34.3% Conservative: 52
Best Local Similarity: 21.6% Mismatches: 160
Query Match: 4.5% Indels: 108
Gaps: 20
DB: 20

US-10-041-030-4 (1-420) x US-10-750-185-38950 (1-3078)
QY 13 AenLysGluProValIySTpLyGlyGluLeuValIleuGlyTyTrHISaNgIyAlaLeuPro 32
Db 1471 AACAGAGAG-----GAGGCGCTCATTTGTTCCAAAGCAGGGAATTTTCATC 1515
QY 33 AaNgIyAaP---ArgGlyAaRgAaLySeR-----ArgPheAlaLeuTyLyS 47
Db 1516 TTAGAGACCTCAAGAAACATTGAGAGAGAGCCCTTGGTGTGATTAATATTCTGCT 1575
QY 48 ArgProLyAaIaAaNgIyValIySProSerThrValHISaIleSerThrProGlnAla 67
Db 1576 AGACCTCAAGCTCTCCAGATTAAAGAAAGAAAGTTG-----GCTGTGTGTATGAAGAT 1635
QY 68 SerLyAlaIleSerCySgIyGlnIleSerIleSerTyTrHISaRgAaNgIn 87
Db 1636 GAAAGACCTTAAAGCTTAAGAGAAAGAAAGTTG-----GCTGTGTGTATGAAGAT 1686
QY 88 ThrValValValGluTyTrHISaPGLyAaPThrAaPhePhe---GlnValGlyAaXg 106
```



```

Db 1687 GCAGTCCCTAAGTTGTTTCATATGTTCCCATGATCCATATGACAACTGTGAAA 1746
Qy 107 SerThrgluserProileasphevalValThraspThrlleserGlyserGlnasnthr 126
Db 1747 TCAACTCCCTTTCTTTTGATTTT-----ACAACCTGCTGAGATGATCAAGATATCT 1800
Qy 127 Asp-----GluAlaglnleThrglnserThrlleserArgphe 139
Db 1801 GAACCTTGATCAAGTTTCAAAATGAAAAGATATCTACACAAAATTAATATATACT 1860
Qy 140 AlAcysArglIleValCysAspArgAsnGluProTyThrlaArgIlePhealaalagly 159
Db 1861 CCATTTTCAG-----GAAAGAAATTCCTCATCTAACAAAGACATCATTTCCCTAGT 1911
Qy 160 PheAspSerSerlyS-----AsnIlePheleuGlyGlnValaalaalyetp 175
Db 1912 ACCAATCTTCAACAAAGACATGTTAACTTATTTTCTAATCAAGATCTGTGCAA--- 1968
Qy 176 LysAsnProAspArglyHsmetAspGlyLeuThThrasnGlyVal-----LeuValMethis 194
Db 1969 AGAAATTTTGTAGG---AAAGACATTAACATGAAAATGAGTACTTATATACGAA 2025
Qy 195 ProArgGlyGlyPheThrglnGlnGlnSerGln-----ProGlyValTPrArgGlnIleSer 212
Db 2026 AACATCTCAAAACATTAACCTCAGAGGGTGAAGATGTCAGATGATGAGATGTAATAGT 2085
Qy 213 ValCysGlyAspValTyThrlleuArgGlnThr-----ArgSerAla 226
Db 2086 ACTGTGTTTATGTCACAGTTTGTGAGATGTTTAAAGAAATATGTCCTTGAACACAG 2145
Qy 227 GlnGlnArgGlyLysLeuValGlnSerGlyThrasnValleuGlnAspGlySerleuIle 246
Db 2146 GAGAGCAATTAACATCACTTCTTAAGAACTATCTCATTTATTTAGACCCCAAAATAGT 2205
Qy 247 AspLeuGlyAlaThrlleuThrArgThrlaAspGlyLeuPheHis----- 263
Db 2206 -----GGTCTACCCACTCTACTCTGAGTCTGATGATCAACATTTATTCA 2256
Qy 264 ---ThrProThrglnLysHsIlleGlnAlaLeuArgGlnGlnIleAsnAlaalaArgPro 282
Db 2257 GAGAGCCCAACAGAAACAAACAAACAAACAAACAAATGATGATATCTAGAGAGAA 2316
Qy 283 GlnCysProValGlyLeuAsnThrlleuAlaPheProSerIleAsnAlaGlyValVal 302
Db 2317 AACTTCAGTTCTGCTGCTGAATAT-----AGTCTCATTTCCACTGACGAAATTTG 2364
Qy 303 GlnGlnLysGlnProTPrAlaTyLeuSerCysGlyHsIlleValHsGlyTyHsAsnTPr 322
Db 2365 AATTATTAAGTTCC----- 2379
Qy 323 GlyHsIleArgSerAspThrglnAlaAsnGlnArgGlnCysProMetCysArgThValGly 342
Db 2380 -----AGATCTATGACTTAC----- 2394
Qy 343 ProTyValProleuThrPheuGlyCysGlnAlaGlyPheTyValaAspAlaGlyProPro 362
Db 2395 -----CCACCA 2400
Qy 363 ThrHsIlePheThrProCysGlyHsIlleValCysSerGlnLysSerAlaTyTyTrpSer 382
Db 2401 ACTTCACACA-----AGTGAATCTTTCACAAATTTCTAGGT 2436
Qy 383 GlnIleProleuProHsIleGlyThr 390
Db 2437 AAATATCTTTAAAGAGGCACT 2460

```

RESULT 6

```

US-10-750-623-38950
; Sequence 38950, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.

```

```

APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MW1100-1
CURRENT APPLICATION NUMBER: US/10/750,623
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 38950
LENGTH: 3078
TYPE: DNA
ORGANISM: Bovine 1986680741510
US-10-750-623-38950

Alignment Scores:
Pred. No.: 34,4 Length: 3078
Score: 102.00 Matches: 88
Percent Similarity: 34.3% Conservative: 52
Best Local Similarity: 21.6% Mismatches: 160
Query Match: 4.5% Indels: 108
Gaps: 20
DB: 8

US-10-041-030-4 (1-420) x US-10-750-623-38950 (1-3078)
Qy 13 AsnLysGlnProValLysTyrgGlyLeuValleuGlyTyraAsnGlyAlaLeuPro 32
Db 1471 AACAGAG-----GAGGCGCTATGTTTCCAAAGACAGGAATTTTCATC 1515
Qy 33 AsnGlyAsp---ArgGlyArgGlySer-----ArgPheAlaLeuTylys 47
Db 1516 TTAGAGACATCAAGAACATTAAGAGAGAGCCCTTGCTGATGATTAATATTCGCT 1575
Qy 48 ArgProLysAlaAsnGlyValLysProSerThValHisValIleSerThProGlnAla 67
Db 1576 AGACCTCAACCTCTCCAGTTTAAAGAAACATTCAGCTGATTCACCTGTCTCAT 1635
Qy 68 SerLysAlaIleSerCysLysGlyGlnHsSerIleSerTyThrlleuSerArgAsnGln 87
Db 1636 GAAAGACTTTAAAGCTTAAGAGAAAGTTCC-----GCTGTGATGAAAGAT 1686
Qy 88 ThrValValValGlyTyThrlaAspLysAspThraspMetPhe---GlnValGlyArg 106
Db 1687 GCAGTCCCTAAGTTGTTTCATATGTTCCCATGATCCATATGACAACTGTGAAA 1746
Qy 107 SerThrgluserProileasphevalValThraspThrlleserGlyserGlnasnthr 126
Db 1747 TCAACTCCCTTTCTTTTGATTTT-----ACAACCTGCTGAGTCAATCATATTTCT 1800
Qy 127 Asp-----GluAlaglnleThrglnserThrlleserArgphe 139
Db 1801 GAACCTTGATCAAGTTTCAAAATGAAAAGATATCTACACAAAATTAATATATACT 1860
Qy 140 AlAcysArglIleValCysAspArgAsnGluProTyThrlaArgIlePhealaalagly 159
Db 1861 CCATTTTCAG-----GAAAGAAATTCCTCATCTAACAAAGACATCATTTCCCTAGT 1911
Qy 160 PheAspSerSerlyS-----AsnIlePheleuGlyGlnValaalaalyetp 175
Db 1912 ACCAATCTTCAACAAAGACATGTTAACTTATTTTCTAATCAAGATCTGTGCAA--- 1968
Qy 176 LysAsnProAspArglyHsmetAspGlyLeuThThrasnGlyVal-----LeuValMethis 194
Db 1969 AGAAATTTTGTAGG---AAAGACATTAACATGAAAATGAGTACTTATATACGAA 2025
Qy 195 ProArgGlyGlyPheThrglnGlnGlnSerGln-----ProGlyValTPrArgGlnIleSer 212
Db 2026 AACATCTCAAAACATTAACCTCAGAGGGTGAAGATGTCAGATGATGAGATGTAATAGT 2085

```

Qy	213	ValCysGluAspValIYThrLeuArgGluThr-----	ArgSerAla	226
Db	2086	ACTGTGTTGATGCCACAGTTTGGAGATGTTTAAAGAAATATGTCCTTGAACCAACGG		2145
Qy	227	GlnGlnArgGlyLeuValGluSerGluThrAsnValLeuGlnAspGlySerLeuIle		246
Db	2146	GAGAGAAATACATCCAGTTCTCTTAAGACCTACTGTCATTATTGACCCCAAAATGTT		2205
Qy	247	AspLeuCyGArgIAlaThrLeuLeuTrpArgThrIlaAspGlyLeuPheHis-----		263
Db	2206	-----GGGTCTCAACCCACTCTACTCCAGCTGCTGATGATCAACAACTATTATTTCAA		2256
Qy	264	---ThrProThrGlnIlyPheHisIleGluIlaLeuArgGlnGluIleAsnAlaIlaArgPro		282
Db	2257	GGAGCCCCACAGAAACAAACAACTAGAGAAATGGTCCAAATTTGATTCTTAGAGAAAGAA		2316
Qy	283	GlnCysProValIglIlyLeuAsnThrLeuAlaPheProSerIleAsnArgIyGluValIAl		302
Db	2217	AAC TTCAGTTCGTGGCTGTAATAT-----AGTCCATTCACACGACAAATTGG		2364
Qy	303	GluGluGlnGlnProTrpAlaTyrlLeuSerCysGlnHisValHisGlyTyrlHisAsnTrp		322
Db	2365	AATTATATAAGTCTCC-----		2379
Qy	323	GlyHisAspSerAspThrGluAlaAsnGluArgGluCysPrometCysArgThrValGly		342
Db	2380	-----AGATCTATGACTTAC-----		2394
Qy	343	ProTyrlValProLeuTrpLeuGlyCysGluAlaGlyPheTyrlValAspAlaGlyProPro		362
Db	2395	-----	CAACCA	2400
Qy	363	ThrHisAlaPheThrProCysGlyHisValCysSerGluIySsrAlaIyETyrlTrpSer		382
Db	2401	ACTCACACA-----AGTGAATCTTCCAAATTTCTAGT		2436
Qy	383	GlnIleProLeuProHisGlyThr		390
Db	2437	AAATACCTTAAAGAGAGCACT		2460
RESULT 7				
US-11-121-086-12/c				
Sequence 12, Application US/11121086				
Publication NO. US20050266459A1				
GENERAL INFORMATION:				
/ APPLICANT: POULSEN, TIM S.				
/ APPLICANT: NIELSEN, KIRSTEN V.				
/ TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES				
/ FILE REFERENCE: 09138, 6000-00000				
/ CURRENT APPLICATION NUMBER: US/11/121,086				
/ CURRENT FILING DATE: 2005-05-04				
/ PRIOR APPLICATION NUMBER: 60/567,570				
/ PRIOR FILING DATE: 2004-05-04				
/ NUMBER OF SEQ ID NOS: 107				
/ SOFTWARE: PatentIn version 3.3				
/ SEQ ID NO 12				
/ LENGTH: 119160				
/ TYPE: DNA				
/ ORGANISM: Homo sapiens				
US-11-121-086-12				
Alignment Scores:				
Pred. No.: 2,47e+03 Length: 119160				
Score: 101.50 Matches: 61				
Percent Similarity: 36.9% Conservative: 22				
Best Local Similarity: 27.1% Mismatches: 98				
Query Match: 4.4% Indels: 45				
DB: 12 Gaps: 13				
US-10-041-030-4 (1-420) x US-11-121-086-12 (1-119160)				
Qy	199	PhenThrGluGluSerGlnProGlyValIleTrpArgGluIleSerValCysGlyAspValIYTr		218

Qy	Db	Sequence	Length	Matches	Conservative	Mismatch	Indels
Qy	Db	15419 TTCAAGGACCTCCCGACGAACTGTTTAACATCATATGCGCTGACATCGTAACAACTGACCC	167891	61	92	28	45
Qy	Db	219 ThrlaunrgsluThrArgseralaglnlnarglylvLeuValGlu-----Ser	167891	61	92	28	45
Qy	Db	15359 TCAGAAAGGAGCTCCGACGAGGCTGTCATGAGGAGCATGCTTCACTGAACTTAACAG	15300	36	98	22	9
Qy	Db	236 GluThrAsnValLeuGlnlnarglyserLeuIleAspLeuCyseGlyValaThrLeuLeuTrp	255	36	98	22	9
Qy	Db	15299 TCTACAGAGTGTCTCCAGCAGGCGAGA---GTAGAACAATGTGGAGCC-----	15255	36	98	22	9
Qy	Db	256 ArgThrAlaLeuArglyLeuPheIleThrProThrGlnIleAlaGlnValLeuArgln	275	36	98	22	9
Qy	Db	15254 -----TTGAGTCCAGCCAAACGAGGTTCAAAATCCAGCTGCGA	15216	36	98	22	9
Qy	Db	276 GluIleAsnAlaAlaArgPro--GlnCyProValGlyLeuAsnThrLeuAlaPhePro	294	36	98	22	9
Qy	Db	15215 CACTTCGACATCGAGCAGCCCTGAACATGCACTTACCTCTCTGAGGCTGTGTTTATG	15156	36	98	22	9
Qy	Db	295 SerIleAsnArglySerValValAlaGlnGlnIleProThrAlaIleLeu---SerCy	313	36	98	22	9
Qy	Db	15155 CATCTCATAGAGG-----GTAGAAACATGTGGCTTGGGAGCTTGGGCTCTGT	15102	36	98	22	9
Qy	Db	314 ---GlyIleValIleGlyThrHis-----AspThrGlyIleAsnSer	326	36	98	22	9
Qy	Db	15101 ACCCCCACTCTCTGACCTTCCACCCACCTTATCCCTTAATATGGATTTGAGAAAC	15042	36	98	22	9
Qy	Db	327 AspThrGlnAlaAsnGlnArgIleCysProMetCysArgThr-----ValGlyProIle	344	36	98	22	9
Qy	Db	15041 GATTCT---TCTAAAGAAAGAGAGACAGACAGAAATGGGAAATGATCAACATCCAGA	14985	36	98	22	9
Qy	Db	345 ValProLeuThrLeuGlnIleCysGlnAlaGlyPheIleVal-----AspAlaGlyPro	361	36	98	22	9
Qy	Db	14984 GTGCTTACCTTGGACGGCTGCTG-TATGATTCAGAGTATCATGCTGAGCCCTGGCT	14926	36	98	22	9
Qy	Db	362 ProThrIleAlaPheThrProCysGlyIleValCysSerGlnIleSerAlaIleIleTrp	381	36	98	22	9
Qy	Db	14925 GGTCCCACTTAAGTCTT-----CACACACATTAATGG	14893	36	98	22	9
Qy	Db	382 -SerGlnIleProLeuProIleGlyThrIleAlaPheIleAlaIleCysProPheCysAl	401	36	98	22	9
Qy	Db	14892 TTCCTCATGTCCCGAGGCCCGAGGAGCCACCTG-----ACAGTGTCCCCCAGCTGAG	14839	36	98	22	9
Qy	Db	401 aThrGlnLeuVal 405	405	36	98	22	9
Qy	Db	14838 GATATGCTTTGTA 14826	14826	36	98	22	9

```

DB: 12 Gap: 13
US-10-041-030-4 (1-420) x US-11-121-086-14 (1-167891)
Oy 199 PheThrgIngluSerGlnProGlyValTPArGluIleSerValCyseGlyAspValTYr 218
Db 166774 TTCAAGTACTCTCCACAGAACGTGTTTAACGATCTATGCTGCTCAACGTAACAGTCAAC 166715
Oy 219 ThrLeuArGluThrArgSerAlaGlnIleArgGlyLysLeuValGlu-----Ser 235
Db 166714 TCAAAAGGACATACCCCAAGAGGTGTCATATGACAGGACGAGCTTCCTCACTAACCAATCTAACAG 166555
Oy 236 GluThrAsnValLeuGlnAspGlySerLeuIleAspLeuCyseGlyAlaThrLeuLeuTP 255
Db 166654 TCTAACACAGGTCTCCACAGAGGACAGG-----GTAAACAATATGTGGACC----- 166610
Oy 256 ArgThrAlaAspGlyLeuPheHisThrProThnGlnLysHisIleGluAlaLeuArgGln 275
Db 166609 -----TTTGAGTCCAGCCCAACAGAGGTTCCAATCCACAGCTGCCA 166571
Oy 276 GluIleAsnAlaAlaArgPro---GlnCyProValGlyLeuAsnThrLeuAlaPhePro 294
Db 166570 CACTTCGAGGCTGGAACGACCTCGAACATGCACTGACCTGCTGATGAGGCTGTGTTTAAg 166511
Oy 295 SerIleAsnArgLysGluValValGluGluLysGlnProTPAlaTyrlen---SerCy 313
Db 166510 CATCTCATAGAGGG-----GTGAAGACATGTGGCTTGGGAGACTTGGGCTCTGT 166457
Oy 314 ---GlyHisValHisGlyTyrlHis-----AsnTPrgIHisArgSer 326
Db 166456 ACCCCCACTCTCTCAGCTTCAACCCACTTTATCCCTTAATATTTGGATTGTGAAC 166397
Oy 327 AsPThrgIuAlaAsnGlnArgGlyCySPrometCyArgThr-----ValGlyProTy 344
Db 166396 GATTCCT---TCTAAAGAAAGAGAGACAGACAGATGCGAAGTAAAGTCAACAGCTCAAG 166340
Oy 345 ValProLeuTPLeuGlyCyseGluAlaGlyPheTyrlVal-----AspAlaGlyPro 361
Db 166339 GTGCCTACCTTGCAGAGGCTGCTG-TATGATTCAGAGTATCATGCTGAGCCCTGGCCCT 166281
Oy 362 ProThrHisAlaPheThrProCyseGlyHisValCyseSerGluLysSerAlaLysTyrlTP 381
Db 166280 GGTCCCACTAAGTCTT-----CACACACTTATTTGG 166248
Oy 382 SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCyseProPheCy 401
Db 166247 TTCCCATGTCCCCACAGGCCAGGCGGACCACTG-----ACAGTGTCCCCACAGTGAG 166194
Oy 401 aThrGlnLeuVal 405
Db 166193 GATATGCTCTTGTGA 166181

RESULT 9
US-11-013-247A-16
; Sequence 16, Application US/11013247A
; Publication No. US2005025496A1
; GENERAL INFORMATION:
; APPLICANT: GAO, ZEREN
; APPLICANT: SHEPPARD, PAUL O.
; APPLICANT: FOX, BRIAN A.
; APPLICANT: HOLLOWAY, JAMES L.
; APPLICANT: JASPERS, STEPHEN R.
; TITLE OF INVENTION: ZENFI2, A TUMOR NECROSIS FACTOR
; FILE REFERENCE: 03-23
; CURRENT APPLICATION NUMBER: US/11/013,247A
; PRIOR FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: 60/530,125
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1664
; TYPE: DNA

```

[illegible]


```

TYPE: DNA
ORGANISM: Mouse
US-11-183-136-41

Alignment Scores:
Pred. No.: 59 Length: 1770
Score: 96.50 Matches: 92
Percent Similarity: 33.2% Conservative: 38
Best Local Similarity: 23.5% Mismatches: 140
Query Match: 4.2% Indels: 122
DB: 12 Gaps: 21

US-10-041-030-4 (1-420) x US-11-183-136-41 (1-1770)

QY 36 ArgGlyArgGlySerArgPheAlaLeuTyrlaArgProlyValaAsnGlyValaLys 55
DB 291 CAGGAGAGATGAGGAGCTGCTACTCTGCAAGAGCT-----329
QY 56 ProSerThrValHisValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGly 75
DB 330 -----CACGTGAGTGGCTTACCCAG-----TCCACTGAGGC 362
QY 76 GlnHis-SerIleSerThrLeuSerArgAsnGlnThr-----ValValValGlnTy 93
DB 363 CACATACACCTCTCATGAAACAAGACCGTGAAGTTACAGACACTGCTGTGACTTT 422
QY 93 rThrHisAspLysAspThrAspMet-----PheGlnValGlyArgSerTh 108
DB 423 TGAGTATGCGCGCCACGGCTCATGTCTCGAAGAGTCCCTGACATGGCGCC--AC 479
QY 108 rGluSerProIleAspPheValValThrAspThrIleSer-----GlySerGlnAsnTh 126
DB 480 CTGGCAGCCCTTACCAGTTCTATGACAGAGAGCTGCATGAGAGCCCTTGGCATGTCTGCCG 539
QY 126 rAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAs 146
DB 540 ACCTGCCCGGAGACATGACCTTCCAGCGCCACCGGCTGCTGCTGC-----585
QY 146 rArgAsnGlnProTyThrAlaArgIlePheAlaAlaGlyPheAspSerSerLysAsnI 166
DB 586 -----ACCGAGAGATACTCACGCG-----TGGGCGAGGGTCCCAAGAAAGACATGT 632
QY 166 e-----PheLysGlyLysAlaAlaLysTrpLysAsnProAsp---GlyHisMetAs 183
DB 633 GCGCTTGAAGTAAAGGAGCGCTTGGCATTTTGGCCGCTTGAACCTGCGTAACTGGA 692
QY 183 rGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThrGlnGln 203
DB 693 CAACCTGTACACAGG-----ATGAGAG 716
QY 203 rGlnProGlyValTrpArgGlnIleSerValCysGlyAspValTyThrLeuArgGlnTh 223
DB 717 CGCAGAGGGCTC-----AAGAGATTCTTCACTTCACTGACTGACT 755
QY 223 rArgSerAlaGlnGlnArg-----GlyLysLeuValGlnSerGlnThr-----237
DB 756 GCGCATGCGCTGCTGCTGCTGCGCTGAGGAGCACTTACCTGAGGAGAACTCTTA 815
QY 238 -----AsnValLeuGlnAspGlySerLeuIleAspLeuCys-GlyAlaThrLeu 254
DB 816 CAAGTACTTTCATGCAATCTTCATATGCAAGATGATGCGAGGATGAGTCAACCTGCA 875
QY 254 eutTrpArgThrAlaAspGlyLeuPheHisThrProThrGlnLysHisIleGlnAlaLeu 274
DB 876 TG-----CCAACCTGTCACAGTGGAG-----898
QY 274 rGlnGlnIleAsnAlaAlaArgProGlnCysProValGlyLeuAsnThrLeuAlaPhe 294
DB 899 -----AGGCGACCTGACAGTGTAGTGTAAACA-----928
QY 294 roserIleAsnArgLysGlnValValGlnGlnLysGlnProTrpAlaTyLeuSerCysG 314
DB 929 -----ACACCAAGGAGCCCACTGTGGAG 953

```

```

QY 314 LysHisValHisGlyTyrlHisAsnTrpGlyHisArgSerAspThrGlnAlaAsnGlnArg 334
DB 954 GT-----GCAAGAAAGAACTCCGCAACAGCGGCTGGGAGC 989
QY 334 LysCysProMetCysArg-----ThrValGlyProTyValPro- 346
DB 990 TGGCTCTTCACTGCTCCGCTGCGCCACCGCTCTCCCATGCTGTGGCGCGGCTCCGC 1049
QY 347 LeuTrpLeuGlyCysGlnAlaGlyPheTyrlValaAspAlaGlyProProThrHisAlaPhe 366
DB 1050 CTTGGCAGTCAACCAAGCAAGCACTACTAT-----GGCCCCCTTGGGAGACAGCTC 1100
QY 367 ThrProCysGlyHisValCysSerGlnLysSerAlaLysTyThrSerGlnIleProLeu 386
DB 1101 CTTCGCGCCCAAGTGTCTCCAGTGCAGAAAGCTGTAGCTATCTGTGCTGTCCCTTC 1160
QY 387 Pro---HisGlyThrHisAlaPheHisAla 395
DB 1161 CCAAGCCAGAGACTTACGCTTTTCAGACT 1190

RESULT 15
US-10-750-185-32846
; Sequence 32846, Application US/10750185
; Publication No. US200502603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENBLUD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: NM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32846
; LENGTH: 1363
; TYPE: DNA
; ORGANISM: Bovine 19866880713026
US-10-750-185-32846

Alignment Scores:
Pred. No.: 67.1 Length: 1363
Score: 94.50 Matches: 58
Percent Similarity: 35.0% Conservative: 28
Best Local Similarity: 23.6% Mismatches: 78
Query Match: 4.1% Indels: 83
DB: 8 Gaps: 13

US-10-041-030-4 (1-420) x US-10-750-185-32846 (1-1363)

QY 205 ProGlyValTrpArgGlu-----210
DB 27 CCGGCGCTTGAAGAGAGAGCTGGGTCTCTGAGATGTCCTCCAGAGAGCTTTGGGAGC 86
QY 211 -----IleSerValCysGlyAspValTyThrLeuArgGlnThrArgSerAlaGln 227
DB 87 CACGTGCTCTCCAGCACTTGGGTGTGCTGTGGGAGCTGAGAGGTAGCCGGGTCCCGCC 146
QY 228 GlnArgGlyLysLeuValGlnSerGlnThrAsnValLeuGlnAspGlySerLeuIleAsp 247
DB 147 TCCAGACCAACCTTACAGGAGCGGCGGCTGAG--CTCTTGTGACAGGCGGTCTCCGGCA 204
QY 248 LeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGlyLeuPheHisThrProThrGln 267
DB 205 TTGTGC-----210

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 2, 2006, 06:06:38 ; Search time 1066 Seconds
(without alignments)
3258.104 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MESPGEHCAPKPERVYK.....ATQVGEQNCIKLIFQGPID 420

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/abs/ABSWEB.spool/US10041030/rnat 01032006 134411 21967/app_query.fasta_1
-DB=Published Applications NA Main -OPMT=fastap -SUFFIX=p2n.rnbp
-NMATCH=0.1 -LOOPT=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR.SCOR=pct
-THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs805h
-USER=US10041030 @CGN 1.1 1026 @rnat 01032006 134411 21967 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA Main:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2290	100.0	1263	6	US-10-085-117-288
2	2290	100.0	5579	5	US-10-197-666A-135
3	2290	100.0	5597	6	US-10-085-117-287
4	2290	100.0	5921	5	US-10-041-030-4
5	2283	99.7	1263	3	US-09-843-905A-7
6	2283	99.7	1263	6	US-10-317-250-7
7	2283	99.7	1263	7	US-10-258-703-7

8	2175.5	95.0	1260	3	US-09-843-905A-5	Sequence 5, App11
9	2175.5	95.0	1260	6	US-10-317-250-5	Sequence 5, App11
10	2175.5	95.0	1260	7	US-10-085-117-285	Sequence 285, App
11	2175.5	95.0	1260	7	US-10-258-703-5	Sequence 5, App11
12	2111.5	92.2	1717	5	US-10-197-666A-133	Sequence 183, App
13	2111.5	92.2	1734	6	US-10-085-117-284	Sequence 284, App
14	1917	83.7	1257	3	US-09-843-905A-3	Sequence 3, App11
15	1917	83.7	1257	6	US-10-317-250-3	Sequence 3, App11
16	1917	83.7	1257	7	US-10-258-703-3	Sequence 3, App11
17	1917	83.7	1257	7	US-10-041-030-1	Sequence 1, App11
18	1917	83.7	1257	5	US-10-443-108-9	Sequence 9, App11
19	1912	83.5	1257	3	US-09-843-905A-1	Sequence 1, App11
20	1912	83.5	1257	6	US-10-317-250-1	Sequence 1, App11
21	1912	83.5	1257	7	US-10-258-703-1	Sequence 1, App11
22	1642	71.7	1338	3	US-09-843-905A-11	Sequence 11, App1
23	1642	71.7	1338	6	US-10-317-250-11	Sequence 11, App1
24	1642	71.7	1338	7	US-10-258-703-11	Sequence 11, App1
25	1482.5	64.7	2508	6	US-10-104-047-1958	Sequence 1958, Ap
26	1302	56.9	2966	10	US-11-097-143-10850	Sequence 10850, A
27	1075	46.9	723	9	US-10-450-763-4485	Sequence 4485, App
28	1064	46.5	202802	6	US-10-085-117-286	Sequence 286, App
29	1048	45.8	155579	6	US-10-085-117-283	Sequence 283, App
30	1022	44.6	541	6	US-10-029-386-22982	Sequence 22982, A
31	996	43.5	31882	10	US-11-097-143-10849	Sequence 10849, A
32	974	42.5	528	6	US-10-029-386-20504	Sequence 20504, A
33	808	35.3	24173	3	US-09-764-891-9284	Sequence 9284, App
34	808	35.3	24173	5	US-10-091-572-855	Sequence 855, App
35	775	33.8	465	3	US-09-867-550-1947	Sequence 1947, App
36	763	33.3	514	6	US-10-029-386-6790	Sequence 6790, App
37	763	33.3	514	6	US-10-029-386-9281	Sequence 9281, App
38	717	31.3	468	3	US-09-918-995-15544	Sequence 15544, A
39	678	29.6	377	3	US-09-867-550-1719	Sequence 1719, App
40	516	22.5	520	9	US-10-756-149-249	Sequence 249, App
41	516	22.5	812	3	US-09-764-891-7046	Sequence 7046, App
42	516	22.5	813	3	US-09-764-891-7044	Sequence 7044, App
43	516	22.5	813	3	US-09-764-891-7047	Sequence 7047, App
44	343	15.0	1587	9	US-10-450-763-4484	Sequence 4484, App
45	330.5	14.4	6014	9	US-10-450-763-13370	Sequence 13370, A

ALIGNMENTS

RESULT 1
US-10-085-117-288
; Sequence 288, Application US/10085117
; Publication No. US2003023334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452004.21
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 288
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-288

Alignment Scores:

Pred. No.: 8.2e-255
Score: 2290.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 6
US-10-041-030-4 (1-420) x US-10-085-117-288 (1-1263)

Qy 1 MetPheSerProGlnGlnGlnGlnHISCyAAlaProAsnLysGluProValLysTrpGly 20
Db 1 ATGTTTTCCCTGGCCGAGGAGAAACAATGGCCCAATAAGAGCCAGTGAATAATACGGG 60
Qy 21 GlnLeuValValLeuGlyTrpAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgLys 40
Db 61 GAGCTGTGTGTGCTCGGGGTCAATGTGTCTTACCAATGAGATGATAGAGACCGAGAGAAA 120
Qy 41 SerArgPheAlaLeuTrpLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 121 AGTAGATTGGCTCTTACAGAGGAGCCCAAGGCAATGTGTCAAAACCAGACCGTTCAT 180
Qy 61 ValLLeSerThrProGlnAlaSerLysAlaLLeSerCyAysGlyGlnHisSerLLeSer 80
Db 181 GTGATATCCAGCCCGAGGATCCAGAGCTATACGCTGCAAAAGGTCAACACAGTATCC 240
Qy 81 TyrThrLeuSerArgAsnGlnThrValValValGlyTrpHisAspLysAspThrAsp 100
Db 241 TACACTTTGTCAAGAAATCAGACTGTGTGTGAGATACACATGATTAAGATACGGAT 300
Qy 101 MetPheGlnValGlyArgSerThrGlnSerProLLeAspPheValLThrAspThrLLe 120
Db 301 ATGTTTCAAGTGTGGCAGATCAACAGAAAGCCCTATCGACTTGTTCACAGACACAT 360
Qy 121 SerGlySerGlnAsnThrAspGluAlaGlnLLeThrGlnSerThrLLeSerArgPheAla 140
Db 361 TCTGGCAGCCAGAAACGAGAGAGCCAGATCCACAGAGACCAATTCAGGTTGGCC 420
Qy 141 CyAArgLLeValCyAAspArgAsnGlnProTrpThrAlaArgLLePheAlaLLeGlyPhe 160
Db 421 TGCAGAGTCTGTGCGACAGAAATGAACCTTACACACAGAGATATTCGCGCGGATTT 480
Qy 161 AspSerSerLysAsnLLePheLeuGlyGlnLysValAlaLysTrpLysAsnProAsnGly 180
Db 481 GACTCTTCCAAAACATATTTCTTGAGAAAAGGACCAAAAGTGGAACCCCGAGCGC 540
Qy 181 HisMetAspGlyLeuThrTrpAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 541 CACATGATGGGCTCACTACTAATGGCGTCTGTGATGATCCACAGGGGGCTTACCC 600
Qy 201 GlnGlnSerGlnProGlyValLTrpArgGlnLLeSerValCyAysLysAspValLTrpThrLeu 220
Db 601 GAGGAGTCCAGCCCGGGGTCTGGCGGAGATCTCTGTGTGAGATGTGTACACCTTG 660
Qy 221 ArgGlnThrArgSerAlaGlnGlnArgLysLeuValGlnSerGlnThrAsnValLeu 240
Db 661 CGAGAAACCAAGTCTGGCCAGCAACAGAGAAAGCTGTGAGAAATGAGACCAAGTCTTG 720
Qy 241 GlnAspGlySerLeuLLeAspLysCyAArgLValAlaThrLeuLeuTrpArgThrAlaAspGly 260
Db 721 CAGGACGGCTCCCTCATTTGACTGTGTGTGGGCCACTCTCTCTGTGAGAAACAGCATGGG 780
Qy 261 LeuPheHisThrProThrGlnLysHisLLeGlnAlaLeuArgGlnLLeAsnAlaAla 280
Db 781 CTTTTCATGACTCCAACTCAGAAAGCAATAGAACCTCCGGCAGGAGATTAAGCGCC 840
Qy 281 ArgProGlnCyAProValLLeLysAsnThrLeuAlaPheProSerLLeAsnArgLysGlu 300
Db 841 CGGCTCAGTGTCTGTGGGCTCAACACCTGGCTTCCCGACATCAACAGAGAAAGAG 900
Qy 301 ValValGlnGlnLysGlnProTrpAlaLTrpLeuSerCyAysGlyHisValHisGlyTyrHis 320
Db 901 GTGTGTGAGAGAGAGAGAGCCCTGGGCAATATCTCAGTTGTGGCCACGTGCACGGGTACAC 960
Qy 321 AsnTrpGlyHisArgSerAspThrGlnAlaAsnGlnArgLysCyAProMetCyAArgThr 340
Db 961 AACTGGGGCCATCGAGTGAACAGGAGCCCAAGAGAGAGGTGTCCCATGTGACAGACT 1020
Qy 341 ValGlyProTrpValLProLeuTrpLeuGlyCyAysGlnAlaGlyPheTrpValAspAlaGly 360
Db 1021 GTGGGCCCCATAGTGTCTGTGCTGTGGCTGTGAGCAGAGATTTTATGAGAGCGCAGGA 1080
Qy 361 ProProThrHisAlaPheThrProCyAArgLysHisValCyAserGlnLysSerAlaLysTyr 380

Db 1081 CCGCAACTCATGCTCTTCACTCCCTGTGAGACGTGTGCTCGGAGAAAGTGTGCAAAATAC 1140
Qy 381 TrpSerGlnLLeProLeuProHisGlyThrHisAlaPheHisAlaAlaCyAProPheCyA 400
Db 1141 TGGTTCAGATCCCGGTGCTCATGGAACCTCAAGATTCACACGCTGTGCTTCTGT 1200
Qy 401 AlaThrGlnLeuValGlyGlnGlnAsnCyAlaLysLeuLLePheGlnGlyProLLeAsp 420
Db 1201 GCTACACAGTGTGTGGGAGCAAAACTGATCAATTAATTTTCCAGGTCCAAATTGAC 1260
RESULT 2
US-10-197-666A-135
; Sequence 135, Application US/1019766A
; Publication No. US2003092037A1
; GENERAL INFORMATION:
; APPLICANT: ASahi KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: Elki phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 135
; LENGTH: 5579
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (177) .. (1436)
US-10-197-666A-135
Alignment Scores:
Pred. No.: 6.8e-254 Length: 5579
Score: 2290.00 Matches: 420
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
US-10-041-030-4 (1-420) x US-10-197-666A-135 (1-5579)
Qy 1 MetPheSerProGlnGlnGlnGlnHISCyAAlaProAsnLysGluProValLysTrpGly 20
Db 177 ATGTTTTCCCTGGCCGAGGAGAAACAATGGCCCAATAAGAGCCAGTGAATAATACGGG 226
Qy 21 GlnLeuValValLeuGlyTrpAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgLys 40
Db 237 GAGCTGTGTGTGCTCGGGGTCAATGTGTCTTACCAATGAGATGATAGAGACCGAGAGAAA 296
Qy 41 SerArgPheAlaLeuTrpLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 297 AGTAGATTGGCTCTTACAGAGGAGCCCAAGGCAATGTGTCAAAACCAGACCGTTCAT 356
Qy 61 ValLLeSerThrProGlnAlaSerLysAlaLLeSerCyAysGlyGlnHisSerLLeSer 80
Db 357 GTGATATCCAGCCCGAGGATCCAGAGCTATACGCTGCAAAAGGTCAACACAGTATTC 416
Qy 81 TyrThrLeuSerArgAsnGlnThrValValValGlyTrpHisAspLysAspThrAsp 100
Db 417 TACACTTTGTCAAGAAATCAGACTGTGTGTGAGATACACATGATTAAGATACGGAT 476

QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValThrAspThrIle 120
 Db 477 ATGTTTCAGGTGGGCAATCAACAGAAAGCCCTATTCACCTTCTGTGCACAGACACAT 536
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 537 TCTGGCAGCCGAAACGAGACGAAAGCCCAATCAACAGACCAATTCAGAGTTGCC 596
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaGlyPhe 160
 Db 597 TGCAGATGCTGTGGCAGAGAAATGAACCTTACACAGACGAAATATTCGCCGCCGATTT 656
 QY 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaIleTyrPheAsnProAspGly 180
 Db 657 GACTCTTCCAAAACATATTTCTTGGAGAAAGGACCAAAAGTGAGAAAAACCCGACGGC 716
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 Db 717 CACATGGATGGGCTCACTACTAATGGCGCTCTGTGATGCATCCAGAGGGGCTTACCC 776
 QY 201 GluGlnSerGlnProGlyValITrPArgGluIleSerValCysGlyAspValTyrThrLeu 220
 Db 777 GAGGAGTCCAGCCCGGGGCTGGCGGAGATCTCTGTGTGAGATGTGTACACCTTG 836
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGlnSerGluThrAsnValLeu 240
 Db 837 CGAGAAACCGAGTGGCGCCGACGAAACGAGAAAGCTGTGAGAAAGTGAGAACACGTCCTG 896
 QY 241 GluAspGlySerLeuIleAspLeuCysGlyValIleThrLeuThrPArgThrAlaAspGly 260
 Db 897 CAGGACCGCTCTCTCATTAACCTGTGTGGGCACTCTCTCTGGAACACGAGAGGG 956
 QY 261 LeuPheHisThrProThrGlnIleHisIleGluAlaLeuArgGlnIleAsnAlaAla 280
 Db 957 CTTTTCATCTCCAACTCAGAGACATAGAGACCTCCGCGACGAGATTACGCCGC 1016
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLeuGln 300
 Db 1017 CGGCTCAGTGTCTGTGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1076
 QY 301 ValValGluGluValGlnProThrAlaTyrLeuSerCysGlyHisValHisIleGlyThrHis 320
 Db 1077 GTGTGGAGGAGAAAGCCCTGGGCAATCTCACTGTGTGCCACGTCGACGGGTACAC 1136
 QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGlnCysProMetCysArgThr 340
 Db 1137 AACCTGGGGCATGGAGATGACAGAGGCCACAGAGGAGAGTCCCATGTGCAGACT 1196
 QY 341 ValGlyProTyrValProLeuTrpLeuGlyCysGlnAlaGlyPheTyrValAspAlaGly 360
 Db 1197 GTGGGGCCCTATGGCTCTCTGTGGCTGTGGAGGACAGATTATTTATGAGACGACAGA 1256
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlyLysSerAlaLysTyr 380
 Db 1257 CCGCAACTATGCTTTCCTCTCTGTGACACCTGTGCTGAGAGAGTCTGCAAAATAC 1316
 QY 381 TrpSerGlnIleProLeuProHisIleGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
 Db 1317 TGGTCTCAAGTCCCGTTGGCTCATGAACTCATGATTTTCAAGCTGTGGCTTCTGTCTGT 1376
 QY 401 AlaThrGlnLeuValGlyGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
 Db 1377 GCTACACAGCTGGTGGGAGCAAAACCTGCATCAATTAATTTTCCAGGTCCCAATTGAC 1436

RESULT 3

US-10-085-117-287
 ; Sequence 287, Application US/10085117
 ; Publication No. US2003023334A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER

FILE REFERENCE: 529452000121
 ; CURRENT APPLICATION NUMBER: US/10/085,117
 ; CURRENT FILING DATE: 2002-02-27
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 361
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 287
 ; LENGTH: 5597
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-085-117-287

Alignment Scores:
 Pred. No.: 6,83e-254 Length: 5597
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-041-030-4 (1-420) x US-10-085-117-287 (1-5597)

QY 1 MetPheSerProGlyGlnGlnIleHisCysAlaProAsnLysGluProValLysTyrGly 20
 Db 177 ATGTTTCCCTGGCCAGAGGAAACCTGGCCCAATTAAGAGCCAGTGAATACGGG 236
 QY 21 GluLeuValIleGluGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgTglys 40
 Db 237 GACTGTGGTGTCTGGGTGACATGTGTCTTACCAATGAGATAGAGAGCGAGAA 296
 QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
 Db 297 AGTAGATTTGCCCTTACAAAGCGCCCAAGGCAAAAGTGTCAAAACCCAGCACCGTCCAT 356
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
 Db 357 GTATATTCACGCCCAAGGATCCAAAGCTATCACTGAGTCAAGGTCAACAGATATTC 416
 QY 81 TyrThrLeuSerArgAsnGlnThrValValIleGluTyrThrHisAspLysAspThrAsp 100
 Db 417 TACACTTGTCAAGATCAGACTGTGTGTGATACACATCATATAGATACGAT 476
 QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
 Db 477 ATGTTTCAGGTGGGCAATCAACAGAAAGCCCTATGACTTCTGTCTCAACAGACTT 536
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 537 TCTGGCAGCCGAAACGAGACGAAAGCCCAATCAACAGACCAATTCAGAGTTGCC 596
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaGlyPhe 160
 Db 597 TGCAGATGCTGTGGCAGAGAAATGAACCTTACACAGACGAAATATTCGCCGCCGATTT 656
 QY 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaIleTyrPheAsnProAspGly 180
 Db 657 GACTCTTCCAAAACATATTTCTTGGAGAAAGGACGAGAAAGTGAGAAAAACCCGACGGC 716
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 Db 717 CACATGATGGGCTCACTACTAATGGCGCTCTGTGATGCATCCAGAGGGGCTTACCC 776
 QY 201 GluGlnSerGlnProGlyValITrPArgGluIleSerValCysGlyAspValTyrThrLeu 220
 Db 777 GAGGAGTCCAGCCCGGGGCTGGCGGAGATCTCTGTGTGAGATGTGTACACCTTG 836
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGlnSerGluThrAsnValLeu 240
 Db 837 CGAGAAACCGAGTGGCGCCGACGAAACGAGAAAGCTGTGAGAAAGTGAGAACACGTCCTG 896
 QY 241 GluAspGlySerLeuIleAspLeuCysGlyValIleThrLeuThrPArgThrAlaAspGly 260

Db 897 CAGACGGCTCCCTCATTTGACCTGTGTGGGCGCACTCTCTCTGGAGAACAGCATGGG 956
Qy LeuphenisthrProthrnglnlyhslleglualaleuargnglnuileanala 280
Db 957 CTTTTCATATCTCCAACTCCAGAACACATAGAACCTTCGGGACGGAGATTAAAGCCGCC 1016
Qy ArgProGlnCysProValGlyLeuasnThrLeuAlaPheProSerIleasnArglysglu 300
Db 1017 CGGGCTCAGTGTCTGTGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1076
Qy 301 ValValGlnGlnlysglnProtrPalaTyrlleuserCysgllyhslvalhslgltYrhis 320
Db 1077 GTGGTGAGAGAGAACGAGCCCTGGGCAATCTCACTGTGGCCACGTCGACGGGTACAC 1136
Qy AsnTrpGlyhslArgSerArpThrGlnAlaAsnGlnuArgGluCysPrometCysArgThr 340
Db 1137 AACTGGGGCCATCGAGATGACGAGAGCCCAAGAGGAGGTCTCCATGTGCAGACT 1196
Qy 341 ValGlyProTyrrValProleuTrpLeuGlyCysGlnuAlaGlyPheTyrrValAspAlaGly 360
Db 1197 GTGGGGCCCTATGTGCTCTGTGGCTTGGCTGTGAGCAGAGATTATTATAGACGAGGA 1256
Qy 361 ProProThrhisAlaPheThrProCysGlyhslValCysSerGlnlySerAlaTyrr 380
Db 1257 CCGCAACTCATGTCTTTCACCTCCCTGTGACACGCTGTGCGAGAGTCTGCAAAATAC 1316
Qy 381 TrpSerGlnIleProleuProHleaglYthrhisAlaPhehisAlaAlaCysProPheCys 400
Db 1317 TGGTTCAGATCCCGTGTGCTCATGAACTCATCTTCCAGCTGCTGCTTGTCTGT 1376
Qy 401 AlaThrGlnleuValGlyGlnuAsnCysIlelyLeuIlePheGlnGlyProIleAsp 420
Db 1377 GCTACACAGCTGGTGGGGAGCAAACTGCATCAATTTTCCAAAGTCCAAATTGAC 1436

RESULT 4

US-10-041-030-3
; Sequence 3, Application US/10041030
; Publication No. US2002150934A1
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Mu, David
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Talarik Inc.
; TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
; FILE REFERENCE: 018761-00681005
; CURRENT APPLICATION NUMBER: US/10/041,030
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/259,502
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5921
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(1763)
; OTHER INFORMATION: human pellino 2
US-10-041-030-3

Alignment Scores:

Pred. No.: 7.4e-254 Length: 5921
Score: 2290.00 Matches: 420
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-041-030-4 (1-420) x US-10-041-030-3 (1-5921)

Qy 1 MetPheSerProGlyGlnGlnGlnhslCysValaProAsnlysgluProVallyTyrrgly 20
Db 501 ATGTTTCCCTGCTGCGAGGAGAAACATGCGGCCCAATTAAGGACGAGTAATACGGG 560
Qy 21 GlnleuValValleuGlyTyrraenglyAlaLeuProAsnGlyAspArgGlyValaGlylys 40
Db 561 GACCTGTGTGTCTGGGTACATAGTGTCTTACCAATGGAATAGAGAGCGAGAGAA 620
Qy 41 SerArgPheAlaLeuTyrrlysaArgProLyAlaAsnGlyVallysbProSerThValHis 60
Db 621 AGTAAATTTGCTCTTACAGCGGCCCCAGGCAATGTGTCAACCCAGACCTTCAT 660
Qy 61 ValIleSerThrProGlnAlaSerlyAlaIleSerCyslysglyGlnhslSerIleSer 80
Db 681 GTATATCCACGGCCAGGATCCAAAGCTATCAGCTGCMAAGTCAACACATATATCC 740
Qy 81 TyrrThrleuserArgaenglnThrValValGlnTyrrThrhisAspIlysaArpThrAsp 100
Db 741 TACACTTTCMAAGGATCAGACTGTGTGTGTGAGTACACATATAGATAGATACGAT 800
Qy 101 MetPheGlnValGlyArgSerThrGlnuserProIleAspPheValValThraArpThrIle 120
Db 801 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGACTTCGTTTCACAGACAGATT 860
Qy 121 SerGlySerGlnaanthraArpGlnuAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 861 TCGGAGCCAGCAACCGAGCGAAAGCCAGATCACACAGAGCAACATATCAGTTCCGC 920
Qy 141 CysArgIleValCysAspArgaenglnProTyrrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 921 TGCAGATCTGTGTGCACAGGAATGAACCTTACACAGCAGGATATTCGCGCGGATT 980
Qy 161 AspSerSerlysaAsnIlePheleuGlyGlylyAlaAlaTyrrTrpIlysaAspProAspGly 180
Db 981 GACTCTTCCAAAAACATATTTCTTGAGAAAGAGCGCAAGATGAAACCCCGACGCG 1040
Qy 181 HisMetAspGlyleuThrThrAsnGlyValleuValMetHisArpArpArglyGlyPheThr 200
Db 1041 CACATGGATGGGTCACTACTATATGGGTCTGTGTATGATGATCAAGAGGGGCTTACC 1100
Qy 201 GlnGlnuserGlnProGlyValITrpaArgGlnIleSerValCysGlyAspValTyrrThrLeu 220
Db 1101 GAGAGATCCAGCCCGGGGTCTGGCGGAGATCTGTGTGTGAGAGATGTGTACACCTTG 1160
Qy 221 ArgGlnThrArgSerAlaGlnGlnuArgGlylyLeuValGlnuserGlnThrAsnValleu 240
Db 1161 CGAAGAACAGGTGCGGCCAGCAACAGAAAGCTGTGAAAGTGAACCAAGCTCTTG 1220
Qy 241 GlnAspGlySerleuIleAspLeuCyrglyAlaThrleuLeuTrpArgThrAlaAspGly 260
Db 1221 CAGGAGGGCTCCCTCATTTGACCTGTGTGGGGCCACTCTCTGTGAGAACAGCAGATGGG 1280
Qy 261 LeuphenisthrProthrnglnlyhslleglualaleuargnglnuileanala 280
Db 1281 CTTTTCATATCTCCAACTCCAGAACACATAGAACCTTCGGGACAGATTAAGCCGCC 1340
Qy 281 ArgProGlnCysProValGlyLeuasnThrLeuAlaPheProSerIleasnArglysglu 300
Db 1341 CGGGCTCAGTGTCTGTGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1400
Qy 301 ValValGlnGlnlysglnProtrPalaTyrlleuserCysgllyhslvalhslgltYrhis 320
Db 1401 GTGGTGAGAGAGAACGAGCCCTGGGCAATCTCACTGTGGCCACGTCGACGGGTACAC 1460
Qy 321 AsnTrpGlyhslArgSerArpThrGlnAlaAsnGlnuArgGluCysPrometCysArgThr 340
Db 1461 AACTGGGGCCATCGAGATGACGAGAGCCCAAGAGGAGGTCTCCATGTGCAGACT 1520
Qy 341 ValGlyProTyrrValProleuTrpLeuGlyCysGlnuAlaGlyPheTyrrValAspAlaGly 360
Db 1521 GTGGGGCCCTATGTGCTCTGTGGCTTGGCTGTGAGCAGAGATTATTATAGACGAGGA 1580
Qy 361 ProProThrhisAlaPheThrProCysGlyhslValCysSerGlnlySerAlaTyrr 380

```

Db      1581 CGCCCAATCATCTTCACTCCCTGTCGACAGCTGTGGAGAAAGTTCGCAAAATAC 1640
Qy      381 TTPSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db      1641 TGGTCTCAGATCCCGCTGCTCATGGAACCATGCAATTCACCGCTGCTTCTGTG 1700
Qy      401 AlaThrGlnLeuValGlyGlnGlnAanCysIleLeuLeuIlePheGlnGlyProIleAsp 420
Db      1701 GCTACACAGCTGGTGGGGAGCAAACTGCATCAATTAATTTTCCAAAGSTCCAAATTGAC 1760

RESULT 5
US-09-843-905A-7
; Sequence 7, Application US/09843905A
; Patent No. US20020168683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-843-905A-7

Alignment Scores:
Pred. No.: 5.32e-254 Length: 1263
Score: 2283.00 Matches: 419
Percent Similarity: 99.8% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 99.7% Indels: 0
DB: 3 Gaps: 0

US-10-041-030-4 (1-420) x US-09-843-905A-7 (1-1263)
Qy      1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAsnIleGluProValIleTyrGly 20
Db      1 ATGTCTTCCCTGCGGAGAGAACTGCGCCCAATAGAGAGCCAGGAAATACGGG 60
Qy      21 GlnLeuValIleuGlyTyrAsnGlyAlaLeuProAsnIleAspArgIleAspArgIle 40
Db      61 GAGCTGTGTGCTCGGGTCAATGTGTCTTACCCATGAGATGAGAGCGAGGAAA 120
Qy      41 SerArgPheAlaLeuTyrIleAspArgIleAspArgIleValIleAspArgIleValIle 60
Db      121 AGTAGATTGCTCCCTTACAGCGGCCCAAGCAATGTGTCAAAACCCACACCGTTCAT 180
Qy      61 ValIleSerThrProGlnAlaSerIleValIleSerCysIleGlnIleSerIleSer 80
Db      181 GTGATATCCACGCCGCCAGGATCCAGAGCTATCAGCTGCAAAAGTCAACAGATATATCC 240
Qy      81 TyrThrLeuSerArgAsnGlnIleThrValIleValIleTyrThrHisAspIleAspArg 100
Db      241 TACACTTTGTTCAGAGATTCAGACTGTGTGTGTGAGTACACATGATTAAGATACGAT 300
Qy      101 MetPheGlnValIleAspArgIleAspArgIleAspArgIleValIleAspArgIle 120
Db      301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCATATGATCTGTTGTCAACAGACGAT 360
Qy      121 SerGlySerGlnAsnThrAspGlnIleGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361 TCTGGCAGCCAGAACAGAGAGCCAGATCAACAGAGACCAATATCCAGGTTCCGCC 420
Qy      141 CysArgIleValIleCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
Db      421 TGCAGATCTGTGTGCAACAGAAATGAACCTTACACAGACGAGATATTCGCCGCCGATTT 480

```

```

Qy      161 AspSerSerIleAsnIlePheLeuGlyGlnValIleAlaIleTyrIleAspArgIle 180
Db      481 GACTCTTCCAAAACATTAATTTCTTGAGATTAAGGACGCAATGTGAAAAACCCCAAGCC 540
Qy      181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db      541 CACATGAGATGGGCTCACTATATGGCGTCTGGATGATCATCCAGAGGGGGCTTCAAC 600
Qy      201 GlnGluSerGlnProGlyValIleTyrArgGlnIleSerValCysGlyAspValIleThrLeu 220
Db      601 GAGAGTCCACAGCCCGGGGTCTGGCGGAGATCTGTCTGTGAGATGTGACACTTGG 660
Qy      221 ArgGlnThrArgSerAlaGlnGlnIleAspArgIleValIleGlnSerGlnThrAsnValLeu 240
Db      661 CAGAAACAGAGTGGGCCAGAACAGAGAAAGCTGTGTGAAGAAGACCAACGCTCTG 720
Qy      241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrArgThrAlaAspGly 260
Db      721 CAGAGCGCTCCCTCATTTGACCTGTGTGGGGCACTCTCTTGAGAAACAGACAGATGGG 780
Qy      261 LeuPheHisThrProThrGlnIleHisIleGlnAlaLeuArgGlnGlnIleAsnAlaIle 280
Db      781 CTTTTCATATCTCAACTCAGAGACATAGAAAGCCCTCCGGCAGAGATTAACGCCGCC 840
Qy      281 ArgProGlnCysProValIleLeuAsnThrLeuAlaPheProSerIleAsnArgIleGlu 300
Db      841 CGGCTCAGATGTCTGTGTGGGGCTCAACACCTGGCTTCCAGACATCAACAGAAAGAG 900
Qy      301 ValIleAlaGlnIleGlyGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyThrHis 320
Db      901 GTGTGTGAGAGAAAGAGAGCCCTGGGCAATATCTCAATGTGTGGCACTGACCGGATCCAC 960
Qy      321 AsnTyrGlyHisArgSerAspThrGlnAlaAsnGlnIleGlnCysProMetCysArgThr 340
Db      961 AACTGGGCGCATTCGAGTACACGAGGCGCAACAGAGAGGAGATGCCCATGTGACAGACT 1020
Qy      341 ValGlyProTyrValProLeuThrLeuGlyCysGlnAlaGlyPheTyrValAspAlaGly 360
Db      1021 GTGGGCCCTATGTGTCTCTGTGGCTGTGGAGGAGATTTATTTATGACAGCAGGA 1080
Qy      361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnIleSerAlaIleTyr 380
Db      1081 CGCCCACTCATGCTTCACTCTCTGTGACACGTGTGTCGGAAGAGCTGCAAAATAC 1140
Qy      381 TTPSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db      1141 TGGTCTCAGATCCCGTGTGCTCATGGAATCATGATTCACGCTGTGGCTTCTGTG 1200
Qy      401 AlaThrGlnLeuValIleGlnGlnAanCysIleLeuLeuIlePheGlnGlyProIleAsp 420
Db      1201 GCTACACAGCTGGTGGGGAGCAAACTGCATCAATTAATTTTCCAAAGSTCCAAATTGAC 1260

RESULT 6
US-10-317-250-7
; Sequence 7, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; CURRENT FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-317-250-7

```


QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 361 TCTGGCAGCGAAGACGAGACGAGAGCCAGATCTCACAGACACCAATATCCAGTTGCGC 420
 QY 141 CysArgIleValCysAspArgAsnGluProThrThrAlaArgIlePheAlaIleGlyPhe 160
 DB 421 TGCAGATCTGTGTCGACAGAAATGAACTTACACACAGAGATTCGCCGCCGATTT 480
 QY 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaIleValIleValIleValIleVal 180
 DB 481 GACTCTTCAAAAACATATTTCTTGAGTAAGGACCAAAAGTGAAGAAACCCCGACGCG 540
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 DB 541 CACATGATGGCTCACTACTAATGCGCTCTGTGTATGATCATCACAGGCGGCTTACCC 600
 QY 201 GluGluSerGlnProGlyValIleTPArgGluIleSerValCysGlyAspValIleThrLeu 220
 DB 601 GAGGAGTCCAGGCCGCGGCTGTGCGGAGATCTCTGTGTGAGATGTGTACACTTG 660
 QY 221 ArgGluThrArgSerAlaGlnGlnArgIleValLeuValGluSerGluThrAsnValLeu 240
 DB 661 CGAAGAACCAAGTGGCGCCAGCAACGAAAGCTGGTGAAGTGAAGCAACGCTCTG 720
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyValIleThrLeuThrThrAlaAspGly 260
 DB 721 CAGACGCGCTCCCTCATTTGACCTGTGTGGGCGCATCTCTCTGTGAAACACGCGATGG 780
 QY 261 LeuPheIleThrProThrGlnIleValIleGluAlaLeuArgGlnIleIleAsnAla 280
 DB 781 CTTTTCATCTCCACTCAGAACACATGAAACCTTCGCGAGGATTAAGCGCC 840
 QY 281 ArgProGlnCysProValIleLeuAsnThrLeuAlaIlePheProSerIleAsnArgIleGlu 300
 DB 841 CGGCTCAGTCTCTGTGGGCTCAACACCTGCTTCCCGACATCAACAGGAAAGAG 900
 QY 301 ValValIleGluGlyGlnProThrAlaIleLeuSerCysGlyValIleValIleGlyThrHis 320
 DB 901 GTGTGTGAGGAGAAAGCAGCCCTGTGGCATATCTCAGTTGTGGCCACGTGCACGCGGTACAC 960
 QY 321 AsnThrGlnIleAspSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
 DB 961 AACTGGGCGCATCGAGTGAACGAGAGCCCAACAGAGGAGGTCTCCATGTGAGGACT 1020
 QY 341 ValGlyProIleValIleProLeuThrLeuGlyCysGluAlaIlePheValIleAspAlaGly 360
 DB 1021 GTGGGCCCTCATGTGCTCTGTGCTGTGCTGTGAGGCAAGATTTTATGACGAGCA 1080
 QY 361 ProProThrIleAlaIlePheThrProCysGlyValIleValIleCysSerGluIleValIleVal 380
 DB 1081 CCGGCAACTCATGTCTTCACTCCCTGTGACACCTGTGCTGTGAGAGTCTGCAAAATAC 1140
 QY 381 ThrSerGlnIleProLeuProIleGlyIleThrHisAlaIlePheHisAlaIleCysProPheCys 400
 DB 1141 TGGCTCAGAACCCGTGCTCTGAGTGAACATCATCTTCACTGTGCTGTGCTTCTGT 1200
 QY 401 AlaThrGlnLeuValIleGluGlnAsnCysIleLeuSerLeuIlePheGlnIleProIleAsp 420
 DB 1201 GCTACACAGCTGCTGTGGGAGCAAACTGCATCAATTAATTTTCCAAAGTCCAAATGAC 1260

RESULT 8
 US-09-843-905A-5
 ; Sequence 5, Application US/09843905A
 ; Patent No. US2002016863A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-A
 ; CURRENT APPLICATION NUMBER: US/09/843, 905A
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,198

; PRIOR FILING DATE: 2000-04-28
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 1260
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; US-09-843-905A-5
 Alignment Scores:
 Pred. No.: 1,56e-241 Length: 1260
 Score: 2175.50 Matches: 399
 Percent Similarity: 97.4% Conservative: 10
 Best Local Similarity: 95.0% Mismatches: 10
 Query Match: 95.0% Indels: 1
 DB: 3 Gaps: 1
 US-10-041-030-4 (1-420) x US-09-843-905A-5 (1-1260)
 QY 1 MetPheSerProGlyGlnGluGluHisCysAlaProAsnIleGluProValIleValIleVal 20
 DB 1 ATGTTTCCCGCGGCGCAGAGGAAACCAGCGCCCAACAGAGAGCGGTGAATACGGG 60
 QY 21 GluLeuValIleGlyIleValIleValIleValIleValIleValIleValIleValIleVal 40
 DB 61 GACTGTGCTGCTGTGGGCTCAATGTGCTTACCTAATGTGACAGGGGAGAGGAA 120
 QY 41 SerArgPheAlaLeuThrIleAspProValIleValIleValIleValIleValIleValIleVal 60
 DB 121 AGCAGATTTGCCCTCTATTAAGGAGCTACGCGAGTGTGTCAACCCAGCAATATCAC 180
 QY 61 ValIleSerThrProGlnAlaSerIleValIleValIleValIleValIleValIleValIleVal 80
 DB 181 ATGCTCTCCACACACAGAGGCTCCAGGCTCAGCTCCAGAGACATCACACATATCG 240
 QY 81 TyrThrLeuSerArgAsnIleThrValIleValIleValIleValIleValIleValIleValIleVal 100
 DB 241 TACACGTGTGACAGGAGCCAGACGTAGTGTGAGATCACACACATTAAGACAGGAC 300
 QY 101 MetPheGlnValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 120
 DB 301 ATGTTTCAGTGGGAGGTGCAACAGAAAGCCCATTTGACTGTGTGTGTGTGTGTGTGTGTGTGT 360
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 361 TCCGGCGGTCAAGCAAGAT--GCCAGATCACACAGGACCAATCTTACGTTGCGCA 417
 QY 141 CysArgIleValCysAspArgAsnGluProIleThrThrAlaArgIlePheAlaIleGlyPhe 160
 DB 418 TGCAGATCTGTGTGACAGAAAGCAAGCATATACAGACGATATTCGGGCGAGATTC 477
 QY 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaIleValIleValIleValIleValIleVal 180
 DB 478 GATTTTCAAAAATATCTTCTTGTGAGGAAAGCAAGCAAAATGAAAAACCTGTATGGA 537
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValIleValIleValIleValIleValIleVal 200
 DB 538 CACATGATGAGTCACTCAATGCTGTCTGTATGATACACCCGAGAGGAGGCTTACACC 597
 QY 201 GluGluSerGlnProGlyValIleTPArgGluIleSerValCysGlyAspValIleThrLeu 220
 DB 598 GAGGAATCCAGCTGAGTGTGAGAGAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 657
 QY 221 ArgGluThrArgSerAlaGlnGlnArgIleValLeuValGluSerGluThrAsnValLeu 240
 DB 658 CGAGAACCAAGTGGCGCCAGCAAGAGGAGAACTGTGAAAGTGAAGCAACGCTCTG 717
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyValIleThrLeuThrThrAlaAspGly 260
 DB 718 CAGACGCGCTCCCTCATTTGACCTGTGTGGGCGCATCTCTGTGAGAAACCGAGATGGC 777
 QY 261 LeuPheIleThrProThrGlnIleValIleGluAlaLeuArgGlnIleIleAsnAlaIle 280

Dh 778 CTTTTCACGCTCTCTACTCAGAACACATGAAAGCCCTCCGGCAGAGATCAATGACGCC 837
Qy 281 ATGTCCTCCACACAGGCGGCTCAAGGCCATGACGATCCAGAGACATCAATATATCG 300
Db 838 CGACCCCAAGTCCCGGCGGCTTAAACCTGGCCCTTCCCGACATCAACGGAGGAA 897
Qy 301 ValValGluGluValGluProTPrAlaTyrLeuSerCysGluValHisGlyTyrHis 320
Db 898 GTGGTGGAAGAAGACGAGCCCTGGGCAATCCTGAGCTGGCCGCAATGACGAGTACAC 957
Qy 321 AsnTPrGluHisArgSerAapThrGluAlaSerGluValGluValGluValGluVal 340
Db 958 AGCTGGGCGCATCGAGCGACCGCGGAAGCCCAACGAGGAGGTCTCCCATGTGACGACT 1017
Qy 341 ValGlyProTyrValProLeuTPrLeuGlyCysGluValGlyPheTyrValAspAlaGly 360
Db 1018 GTGGGCCCCCTACGTCCTCTCTGAGCTGTGAGGAGGATTTATATGTCGAGCGGA 1077
Qy 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluValSerAlaValTyr 380
Db 1078 CCCCACACTCAGCGCTTTCACCCCTGGGGCAGCTGTTCAGAAAGTCTGCCAATGAC 1137
Qy 381 TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1138 TGGTCGAGATCCCACTGCGCCACGGAACGCAACGCTTTCATGCGCGCTGTCCGTTCTGC 1197
Qy 401 AlaThrGlnLeuValGlyGluGlnAnCysIleValLeuIlePheGlnGlyProIleAsp 420
Db 1198 GCCACGAGCTGTTGGTGAACAGAACTGCATCAAAATTGATTTTCCAAAGTCCAGTGAC 1257

RESULT 9

US-10-317-250-5
; Sequence 5, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2890-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; CURRENT FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-317-250-5

Alignment Scores:

Pred. No.: 1.56e-241 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.4% Conservation: 10
Best Local Similarity: 95.0% Mismatches: 10
Query Match: 95.0% Indels: 1
Gaps: 1

US-10-041-030-4 (1-420) x US-10-317-250-5 (1-1260)

Qy 1 MetPheSerProGlyGlnGluGluHisCysAlaProSerLeuValProValTyrGly 20
Db 1 ATGTTTTCCTCCGCGGCGGAGGAAACCAAGCCCTCCCAAGAGAGCCGGTGAATTAAGGG 60
Qy 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProSerGlyAapArgGly 40
Db 61 GAGCTGTGTCTCGGGGTACAAATGTGCTTAACTTAATGTGACAGGGGAGAGGAAA 120
Qy 41 SerArgPheAlaLeuTyrValArgProGlyAlaSerGlyValValProSerThrValHis 60
Db 121 AGCAGATTTGCTCTTAATAGCGGACCTTAACGCTGTGTGCAAAACCAAGACATATCAC 180
Qy 61 ValIleSerThrProGlnAlaSerTyrAlaIleSerCysValGlyGlnHisSerIleSer 80

Db 181 ATGTCCTCCACACAGGCGGCTCAAGGCCATGACGATCCAGAGACATCAATATATCG 240
Qy 81 TyrThrLeuSerArgSerGlnThrValValGluTyrThrHisAlaPheThrAsp 100
Db 241 TACACGTTTCAAGGAGCCAGAGGTAGTGTGAATACACACATTAAGACAGGAC 300
Qy 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
Db 301 ATGTTTCAGTGGGCGAGTCAACGAAGGCCCATTTGATCTGTGGTCAACGACGGTT 360
Qy 121 SerGlySerGlnAnThrAapGluValGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCCGGCGGTCAAGACAAAGAT--GCCAGATACACAGAGACCATCTCTAGTTTCGCA 417
Qy 141 CysArgIleValCysAapArgSerGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 418 TGCAGATCTGTGTACAGGAACGAGCCATTAACGACGCAATATTCGGGAGGATTC 477
Qy 161 AspSerSerTyrAsnIlePheLeuGlyGluValAlaAlaValTPrPheAsnProAspGly 180
Db 478 GATTCCTCCAAAATATCTTTCTTGAAGAGAAAGCAGCAAAATGGAATAAATCTGATGA 537
Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 538 CACATGATGATGACTACTACATGATGTGTCTAGTATGACACCGCAGAGGAGGCTTCACC 597
Qy 201 GluGluSerGlnProGlyValTPrArgGluIleSerValCysGlyValPheValTyrThrLeu 220
Db 598 GAGGAATCCAGGCTGTGAAGTCTGGAAGAGATCTGTCTGTGGAGATGTGACACCTTG 657
Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnValLeu 240
Db 658 CGAGACACAGTCTCGGCCGACGAGGAGGAACTGTGAAGTGAAGACCAACGTCCTG 717
Qy 241 GlnAspGlySerLeuIleAapLeuCysGlyValThrLeuLeuTPrArgThrAlaAspGly 260
Db 718 CAAGACGGCTCCTCATTTGACCTGTGTGGGGCCACTCTCTCTGGAAGACCGCAGATGCG 777
Qy 261 LeuPheHisThrProThrGlnValHisIleGluValLeuArgGlnGluIleAsnAlaAla 280
Db 778 CTTTTCACGCTCTCTCTCAGAAACATTAAGAACCTCCGGCAGAGATCAATGACGCC 837
Qy 281 ArgProGlnCysProValGlyLeuAnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
Db 838 CGACCCCAAGTCCCGTGGGCTTAAACCTTGCTTCCCGACATCAACCGGAAGAA 897
Qy 301 ValValGluGluValGluProTPrAlaTyrLeuSerCysGlyValHisGlyTyrHis 320
Db 898 GTGGTGGAAGAAGACGAGCCCTGGGCAATCCTGAGCTGGGCGCATGACGAGTACAC 957
Qy 321 AsnTPrGluHisArgSerAapThrGluAlaSerGluValGluValGluValGluVal 340
Db 958 AGCTGGGCGCATCGAGCGACCGCGGAAGCCCAACGAGGAGGTCTCCCATGTGACGACT 1017
Qy 341 ValGlyProTyrValProLeuTPrLeuGlyCysGluValGlyPheTyrValAspAlaGly 360
Db 1018 GTGGGCCCCCTACGTCCTCTCTGAGCTGTGAGGAGGATTTATATGTCGAGCGGA 1077
Qy 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluValSerAlaValTyr 380
Db 1078 CCCCACACTCAGCGCTTTCACCCCTGGGGCAGCTGTTCAGAAAGTCTGCCAATGAC 1137
Qy 381 TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1138 TGGTCGAGATCCCACTGCGCCACGGAACGCAACGCTTTCATGCTCCGCTGTCTGC 1197
Qy 401 AlaThrGlnLeuValGlyGluGlnAnCysIleValLeuIlePheGlnGlyProIleAsp 420
Db 1198 GCCACGAGCTGTTGGTGAACAGAACTGCATCAAAATTGATTTTCCAAAGTCCAGTGAC 1257

RESULT 10

US-10-085-117-285

Sequence 285, Application US/10085117
 Publication No. US2003023234A1
 GENERAL INFORMATION:
 APPLICANT: Morris, David W.
 APPLICANT: Engelhard, Eric K.
 TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
 FILE REFERENCE: 52945200121
 CURRENT APPLICATION NUMBER: US/10/085,117
 CURRENT FILING DATE: 2002-02-27
 PRIOR APPLICATION NUMBER: US 09/798,586
 PRIOR FILING DATE: 2001-03-02
 NUMBER OF SEQ ID NOS: 361
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 285
 LENGTH: 1260
 TYPE: DNA
 ORGANISM: Mus musculus
 US-10-085-117-285

Alignment Scores:
 Pred. No.: 1,566-241 Length: 1260
 Score: 2175.50 Matches: 399
 Percent Similarity: 97.4% Conservative: 10
 Best Local Similarity: 95.0% Mismatches: 10
 Query Match: 95.0% Indels: 1
 DB: 6 Gaps: 1

US-10-041-030-4 (1-420) x US-10-085-117-285 (1-1260)

QY 1 MetPheSerProGlyGlnGluGluHisCysAlaProAsnLysGluProValLysTyrGly 20
 DB 1 ATGTTTCCCGGGCCGAGGAGAACCCAGCGCCCAACAGAGCCGGTGAATACGG 60
 QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgLys 40
 DB 61 GACGTGGTGTCTCGGGGTCAATAGTGCTTACCTAATGTGACAGGGCAGAGAGAA 120
 QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
 DB 121 AGCAGATTGCTCCCTAATAGCGGACCTACGCGGTGTCAAAACCCAGACAAATCCAC 180
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
 DB 181 ATGCTCTCCACACACAGCGGCTCCAGCCATGCTCCAGAGACATCCACAGCATATCG 240
 QY 81 TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAspLysAspThrAsp 100
 DB 241 TACACGTTGTCACAGAGCCAGCGTAGTGAGTACACACACGATTAAGACACGAC 360
 QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
 DB 301 ATGTTTCAGGTGGGAGGTCAACGAAAGCCCATTCCTCGGTGCACAGACACGGTT 360
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 361 TCCGGCGGTGAGAAAGAT--GCCCATGTCACACAGACACCATCTTACGTTGCA 417
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 DB 418 TGCAGATCTCTGTGACAGAAAGACCATATACAGACGCAATTCGGCGAGATTC 477
 QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTyrLysAsnProAspGly 180
 DB 478 GATTCCTCCAAAATATATCTTCTTGAGAGAAAGACAAATGGAATAACCTGATGGA 537
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProAspArgLysIlePheThr 200
 DB 538 CACATGATGACTACTACCAATGCTCTAGTAGACCCGCAAGAGGCTTCACC 597
 QY 201 GluGluSerGlnProGlyValTyrParGluIleSerValCysGlyAspValTyrThrLeu 220
 DB 598 GAGGAAATCCAGCGCTGAGTCTGGAGAGAAATCTCTCTGTGGGATGTGTACACCTTG 657

QY 221 ArgLysThrArgSerAlaGlnGluArgLysLeuValGluSerGlnThrAsnValLeu 240
 DB 658 CAGAGACCAAGCTCGGCCACGACGAGGGAGAACTGTGGAAATGAGACCAACGCTCTG 717
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrParGlnAlaAspGly 260
 DB 718 CAAGACGGCTCCCTCATTTACCTGTGTGGGGCACTCTCTGTGAGAACCGGAGATGGC 777
 QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnIleAlaAla 280
 DB 778 CTTTTCAGCTCTCTACTCAGAGCAATAGAGCCCTCCGCGAGAGATCAATGACGCC 837
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
 DB 838 CAGCCCAAGTCCCTGAGGCTTTACACCTTGCTGCTTCCCAAGTCAACCGAAGAA 897
 QY 301 ValValGluGluLysGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
 DB 898 GTGTGGAAAGAGAGAGCCCTGGGCTATGAGCTGGGCGCATGTGACCGCTACAC 957
 QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgLysCysProMetCysArgThr 340
 DB 958 AGCTGGGGCCATCGAGCGACCGGAAAGCCAAAGAGGAGAGTGTCCATGTGCAGAGACT 1017
 QY 341 ValGlyProTyrValProLeuTyrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
 DB 1018 GTGGGCCCTTACCTCTCTGTGGGCTGTGAGGCGAGATTATATGATGCGGGA 1077
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyr 380
 DB 1078 CCCCACACTACGCTTACACCCCTGGGCAAGCTGTGTTCAGAAAGCTGCGCAAGTAC 1137
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
 DB 1138 TGTGTCCAGATCCCACTGCCCCACGAGACGACGCTTCAATGCCCGCTGTCCGTGCG 1197
 QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnIleProIleAsp 420
 DB 1198 GCCACGACGCTGTTGTGAAACGAACTGATCAATTAATTTCCAAAGGTCCAGTGAC 1257

RESULT 11
 US-10-258-703-5
 Sequence 5, Application US/10258703
 Publication No. US2004003419A1
 GENERAL INFORMATION:
 APPLICANT: Bird, Timothy A.
 APPLICANT: Cosman, David J.
 TITLE OF INVENTION: HUMAN BELLINO POLYPEPTIDES
 FILE REFERENCE: 2990-US
 CURRENT APPLICATION NUMBER: US/10/258,703
 CURRENT FILING DATE: 2002-10-24
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 5
 LENGTH: 1260
 TYPE: DNA
 ORGANISM: Mus musculus
 US-10-258-703-5

Alignment Scores:
 Pred. No.: 1,566-241 Length: 1260
 Score: 2175.50 Matches: 399
 Percent Similarity: 97.4% Conservative: 10
 Best Local Similarity: 95.0% Mismatches: 10
 Query Match: 95.0% Indels: 1
 DB: 7 Gaps: 1

US-10-041-030-4 (1-420) x US-10-258-703-5 (1-1260)

QY 1 MetPheSerProGlyGlnGluGluHisCysAlaProAsnLysGluProValLysTyrGly 20
 DB 1 ATGTTTCCCGGGCCGAGGAGAACCCAGCGCCCAACAGAGCCGGTGAATACGG 60

```

Qy 21 GluLeuValIleuGlyTYrAsnGlyValAlaLeuProAsnGlyAspArgGly 40
Db 61 GAGTGGTGGTCCGGGGTACAATGGTCTTACTTAATGGTGAACAAGGAGGAGAA 120
Qy 41 SerArgPheAlaLeuTYrLYsArgProLYsAlaAsnGlyValLYsProSerThrValHis 60
Db 121 AGCAGATTTCCTCTTAAGCCGACCTACGCCGTGGTGTCAAAACCCAGACAAATCCAC 180
Qy 61 ValIleSerThrProGlnAlaSerLYsAlaIleSerCysLYsGlnHisSerIleSer 80
Db 181 ATGGTCTCCACACACAGCGCTCCAGAGCCATCTCCAGAGCAATCCACGCAATTCG 240
Qy 81 TYrThrLeuSerArgAsnGlnThrValValGluTYrThrHisAspLYsAspThrAsp 100
Db 241 TACACGTTGTCCACGAGCCAGACGGTAGTGGTAGATACACACGATTAAGAACAAGAC 300
Qy 101 MetPheGlnValGlyAspSerThrGluSerProIleAspPheValValThrAspThrIle 120
Db 301 ATGTTTCAGGTGGGCGAGGTCAACAGAAAGCCCAATTGACTTGTGTACAGACACGCTT 360
Qy 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCCGGCGGTCAAGAACGAAAT--GCCAGATTCACACAGACACCATCTTAAAGTTGCA 417
Qy 141 CysArgIleValCysAspArgAsnGluProTYrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 418 TGCAGGATCGTGTGTGACAGGAACGAGCANAATACAGACGCAATTCGGCGGAGATTC 477
Qy 161 AspSerSerLYsAsnIlePheLeuGlyGluLYsAlaAlaLYsTYrLYsAsnProAspGly 180
Db 478 GATTCTTCAAAAATACTTCTTCTTGAGAGAAACAGCAAAATGTGAAAACCCCTGATGGA 537
Qy 181 HisMetAspGlyLeuThrThrAsnGlyValIleuValMetHisProAspGlyGlyPheThr 200
Db 538 CACATGGATGACTCACTACCAATGGTGTCTTAGTGAAGACCCGCAAGAGGGCTTCAAC 597
Qy 201 GluGluSerGlnProGlyValTYrArgGluIleSerValCysGlyAspValTYrThrLeu 220
Db 598 GAGGAATCCCAAGCCTGAGAGTGTGAGAGAGATCTCTGTGTGGGATGTGTACACTTG 657
Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyLYsLeuValGluSerGluThrAsnValLeu 240
Db 658 CGAAGACCAAGGTGGGCGCCAGCAAGGGGAAAGCTGGGGAAGATGACCAACAGTCTCG 717
Qy 241 GluAspGlySerLeuIleAspLeuCysGlyValaThrLeuLeuTPArgThrAlaAsnGly 260
Db 718 CAAGACGCTCTCCATTGACCTGTGGGGCCACTCTCTCTGAGAAACCGCAGATGGC 777
Qy 261 LeuPheHisThrProThrGlnLYsHisIleGlnAlaLeuArgGlnGlnIleAsnAlaAla 280
Db 778 CTTTTCACGCTCTTACTCAGAACCACTAAGAACCCCTCCGGCAGAGATCAATGCAGCC 837
Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLYsGlu 300
Db 838 CGAACCCAGTGCCTCCGTGGGCTTAAACACCTGGCCTTCCCAACATCAACCGGAGGAA 897
Qy 301 ValValGluGluLYsArgInProTPAlaTYrLeuSerCysGlyHisValHisGlyTYrHis 320
Db 898 GTGGTGAAGAAGAACGCTCGGCAATACCTGAGCTGGCCACTGTGCAGCGCTACAC 957
Qy 321 AsnTPArgLYsAspSerThrGlnAlaAsnGlnArgGluCysProMetCysArgThr 340
Db 958 AGCTGGGGCCATCCGAGGACGCGGAAACCAAGAGAGGTCTCCATGTGCAGACT 1017
Qy 341 ValGlyProTYrValProLeuTPLeuGlyCysGlnAlaGlyPheTYrValAspAlaGly 360
Db 1018 GTGGGCCCCCTACGCTCTCTGTGGCTGTGAGGAGGAGATTTAATGATGCGAGGA 1077
Qy 361 ProProThrHisAlaPheThrProCysGlyHisIleValCysSerGlnLYsSerAlaLYsTYr 380
Db 1078 CCCCACACTCAAGCTTTCACCCCTCGCGGACGCTGTTCAGAAAATGTGCGCAAGTAC 1137
Qy 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400

```

```

Db 1138 TGGTCGAGATCCCACTGCCCAAGGACGACGCGTTTCAATGCCCTGTCCGTTCTGC 1197
Qy 401 AlaThrGluLeuValGlyGluGlnAsnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
Db 1198 GCCACGACGTGGTGGTGAACAGAACTGCATCAAAATGATTTTCCAAAGTCCAGTGGAC 1257

RESULT 12
US-10-197-666A-133
; Sequence 133, Application US/10197666A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAH I KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: E1k1 phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197, 666A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 133
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (218) .. (1474)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1288..1301
; OTHER INFORMATION: n=A or C or G or T
US-10-197-666A-133

Alignment Scores:
Pred. No.: 6,41e-234 Length: 1717
Score: 2111.50 Matches: 388
Percent Similarity: 95.2% Conservative: 12
Best Local Similarity: 92.4% Mismatches: 19
Query Match: 92.2% Indels: 1
DB: Gaps: 1

US-10-041-030-4 (1-420) x US-10-197-666A-133 (1-1717)
Qy 1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAsnLYsGluProValLYsTYrGly 20
Db 218 ATGTTTCCCGGGCCAGAGGAACCAAGCGCCCAACAAAGAGCGCGTGAATACAGG 277
Qy 21 GluLeuValIleuGlyTYrAsnGlyValAlaLeuProAsnGlyAspArgGlyAspGly 40
Db 278 GAGTGGTGGTCCCTTAAGCGGTACAATGGTGTCTTACTTAATGGTGAACAAGGAGGAGAA 337
Qy 41 SerArgPheAlaLeuTYrLYsArgProLYsAlaAsnGlyValLYsProSerThrValHis 60
Db 338 AGCAGATTTCCTCTTAAGCCGACCTACGCCAGTGGTGTCAAAACCCAGACAAATCCAC 397
Qy 61 ValIleSerThrProGlnAlaSerLYsAlaIleSerCysLYsGlnHisSerIleSer 80
Db 398 ATGGTCTCCACACACAGCGCTCCAGAGCCATACGCTCCAGAGGACATCAACGATATCG 457
Qy 81 TYrThrLeuSerArgAsnGlnThrValValGluTYrThrHisAspLYsAspThrAsp 100
Db 458 TACACGTTGTCCACGAGCCAGACGGTAGTGGTAGATACACACGATTAAGAACAAGAC 517

```

```

QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValThrAspThrIle 120
Db 518 ATGTTTCAGGTGGGCGAGTCAACAGAAAGCCCACTTACCTTCGGGTCACAGACGGTT 577
QY 121 SerGlySerGlnAsnThrAspGluAglInIleThrGlnSerThrIleSerArgPheAla 140
Db 578 TCCGGGGGTTCAGAACGAAGAT---GCCAGATCAACAGAGCAACATCTCTAGGTTCCGA 634
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAglIlePhe 160
Db 635 TGCAGATCCGTGTGTGACAGAAACAGCCATACACAGCATATTCGGGGCAGGATTC 694
QY 161 AspSerSerIleValAsnIlePheLeuGlyValValAlaIleValTyrPheAsnProAspGly 180
Db 695 GATTCCTCCAAAATATCTTCTTGTGGAGAAAGACCAAAATGGAAGAAACCTCGATGGA 754
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgIleGlyPheThr 200
Db 755 CACATGATGATGACTCACTACCAACGGTGTCTGTGTATGACCCGCAAGAGGCTTCACC 814
QY 201 GlnGlnSerGlnProGlyValITPArgGluIleSerValCysGlyAspValITyrThrLeu 220
Db 815 GAGGAATCCAGCCTGAGTCTGGAGAGAGATCTCTGTCTGTGGGAGTATATACACCTTG 874
QY 221 ArgGluThrArgSerArgIleGlnGlnValGlyValLeuValGlnSerGluThrAsnValLeu 240
Db 875 CGAGAGACCGAGTGGCCGAGAGAGAGGGAAGCTGTGTGAAAGTGAACCAACGTCCTG 934
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyValIThrLeuLeuTyrPArgThrAlaAspGly 260
Db 935 CAAGACGGCTCCTCATTTGACTGTGTGTGGGGCCACTCTCTCTGGAGAACCGCAGATGGC 994
QY 261 LeuPheHisThrProThrGlnIleValHisIleGluAlaLeuArgGlnIleAsnAlaIle 280
Db 995 CTTTTCACGCTCCTACTCAGAAAGCAGATAGAAAGCCTCCGGCAGAGATCAATGACGCC 1054
QY 281 ArgProGlnCysProValIleValAsnThrLeuAlaPheProSerIleAsnArgValGlu 300
Db 1055 CGAGCCCGAGGCCCCGGGGGCTTMAACCCCTGGCCTTCCCGAGCATCAACCGGAAGGAA 1114
QY 301 ValValGlnGlnIleValGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
Db 1115 GTGGTGAAGAAAGACGCCCCCTGGGCACTACCTGAGCTGGCGCATGTGCACGGCTTACAC 1174
QY 321 AsnTyrGlyHisArgSerAspThrGlnAlaAsnGluArgIleCysPrometCysArgThr 340
Db 1175 CACTGGGGCCATCGAGCGACAGCGAAGCCACAGAGGAGGAGTGTCCCATGTGCATGACT 1234
QY 341 ValGlyProTyrValProLeuTyrPheGlyCysGlnAlaGlyPheTyrValAspAlaGly 360
Db 1235 GTGGGGCCCCACGCTCTCTGTGGCTGGGCTGTGAGGACAGATTATTTATGCGANNNNNN 1294
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluYssSerAlaValTyr 380
Db 1295 NNNNNNNCTATGTTTTCACCCCTGGGGGACGCTGTTCAGAGAGTCTGCCAAGTAC 1354
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIaCysProPheCys 400
Db 1355 TGGTTCGAGATCCACATCGCCCAAGCAACCAACCGCTTTCACGCGGCTGTCCGTTCTGC 1414
QY 401 AlaThrGlnLeuValGlyGlnIleAsnCysIleValLeuIlePheGlnGlyProIleAsp 420
Db 1415 GCCACGACGCTGGTGTGTGAACAGAACTGATCAATTTGATTTCCAAAGTCCAGTGAC 1474

```

```

FILE REFERENCE: 529452000121
CURRENT APPLICATION NUMBER: US/10/085,117
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 361
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 284
LENGTH: 1734
TYPE: DNA
ORGANISM: Mus musculus
FEATURES:
NAME/KEY: variation
LOCATION: (1)-(1734)
OTHER INFORMATION: n = any nucleotide
US-10-085-117-284

Alignment Scores:
Pred. No.: 6.5e-234 Length: 1734
Score: 211.50 Matches: 388
Percent Similarity: 95.2% Conservative: 12
Best Local Similarity: 92.4% Mismatches: 19
Query Match: 92.2% Indels: 1
DB: Gaps: 1

US-10-041-030-4 (1-420) x US-10-085-117-284 (1-1734)
QY 1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAsnGluProValITyrGly 20
Db 218 ATGTTTCCCGGGCGAGAGAAACCAAGCGCCCAACAGAGACCGGTGAATACAGG 277
QY 21 GlnLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyAspArgGly 40
Db 278 GACCTGTGTCTCTGGGAGTCAATGTGTCTTAACTGAATGTGACAGGGGCAAGAAAGAA 337
QY 41 SerArgPheAlaLeuTyrIleValArgProValAlaAsnGlyValITyrProSerThrValHis 60
Db 338 AGCAGATTTACCTCTTATTAAGCGGAGCTACGCAAGTGTGTCAAAACCGACACATCCAC 397
QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCysIleGlyGlnHisSerIleSer 80
Db 398 ATGCTTCCACACACAGAGGGGTCCAGGCGCATGAGTCCAGAGACATCAACGATATCG 457
QY 81 TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAlaPheAspThrAsp 100
Db 458 TACACGTTTGCACGAGCGCCAGCGGTAGTGTGTGAGTACACACACATTAAGACACAGAC 517
QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValIThrAspThrIle 120
Db 518 ATGTTTCAGGTGGGCGAGTCAACAGAAAGCCCACTTACCTTCGGTCAACAGACGGTT 577
QY 121 SerGlySerGlnAsnThrAspGluAglInIleThrGlnSerThrIleSerArgPheAla 140
Db 578 TCCGGGGGTTCAGAACGAAGT---GCCAGATCAACAGAGCAACATCTCTAGGTTCCGA 634
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAglIlePhe 160
Db 635 TGCAGATCCGTGTGTGACAGAAACAGCCATACACAGCATATTCGGGGCAGGATTC 694
QY 161 AspSerSerIleValAsnIlePheLeuGlyValValAlaIleValTyrPheAsnProAspGly 180
Db 695 GATTCCTCCAAAATATCTTCTTGTGGAGAAAGACCAAAATGGAAGAAACCTCGATGGA 754
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgIleGlyPheThr 200
Db 755 CACATGATGATGACTCACTACCAACGGTGTCTGTGTATGACCCGCAAGAGGCTTCACC 814
QY 201 GlnGlnSerGlnProGlyValITPArgGluIleSerValCysGlyAspValITyrThrLeu 220
Db 815 GAGGAATCCAGCCTGAGTCTGGAGAGAGATCTGTGTGTGGGAGTATATACACCTTG 874
QY 221 ArgGluThrArgSerArgIleGlnGlnValGlyValLeuValGlnSerGluThrAsnValLeu 240

```


This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 2, 2006, 05:56:29 ; Search time 253 Seconds
(without alignments)
2950.893 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MESPGEHNCAPKPEPVKVG.....ATQVGEQNCIKIFQGPID 420

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q/abses/ABSSWEB.spool/US10041030/runat 01032006 134408 21912/app query.fasta_1
-DB=Issued Patents NA -QFMT=fasta -SUFFIX=p2n.rml -MINMATCH=0.1 -LOOPTL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cti
-LIST=45 -DOCALL=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=abses -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abses04 -USER=US10041030 @CEN 1 1 193 @runat 01032006 134408 21912
-NCPU=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARM.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/H COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCUS COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2283	99.7	1263	3	US-09-843-905A-7
2	2175.5	95.0	1260	3	US-09-843-905A-5
3	1917	83.7	1257	3	US-09-843-905A-3
4	1912	83.5	1257	3	US-09-843-905A-1
5	1642	71.7	1338	3	US-09-843-905A-11
6	1482.5	64.7	2508	3	US-10-104-047-1958
7	120.5	5.3	2750	2	US-08-136-277-1
8	120.5	5.3	2750	2	US-08-479-403-1
9	120.5	5.3	2750	3	US-08-835-734-1

C 10	113.5	5.0	4403765	3	US-09-103-840A-2	Sequence 2, Appl
C 11	113.5	5.0	4411529	3	US-09-103-840A-1	Sequence 1, Appl
C 12	109	4.8	1580	3	US-09-252-991A-14397	Sequence 14397, A
C 13	109	4.8	3801	3	US-09-252-991A-14280	Sequence 14280, A
C 14	107	4.7	86414	3	US-09-949-016-12345	Sequence 12345, A
C 15	107	4.7	86414	3	US-09-949-016-12345	Sequence 12345, A
C 16	106.5	4.7	2240	3	US-09-799-451-571	Sequence 571, App
C 17	106	4.6	1314	3	US-09-902-540-2044	Sequence 2044, App
C 18	106	4.6	2574	3	US-09-902-540-4745	Sequence 4745, App
C 19	106	4.6	25254	3	US-09-902-540-1233	Sequence 1233, App
C 20	106	4.6	4403765	3	US-09-103-840A-1	Sequence 1, Appl
C 21	106	4.6	4411529	3	US-09-103-840A-2	Sequence 2, Appl
C 22	105	4.6	6476	3	US-09-854-133-428	Sequence 428, App
C 23	105	4.6	22512	3	US-09-902-540-1220	Sequence 1220, App
C 24	105	4.6	34030	3	US-09-949-016-12248	Sequence 12248, A
C 25	105	4.6	34031	3	US-09-949-016-13540	Sequence 13540, A
C 26	104	4.5	2133	3	US-09-252-991A-9129	Sequence 9129, App
C 27	104	4.5	44120	3	US-09-949-016-14151	Sequence 14151, A
C 28	104	4.5	44120	3	US-09-949-016-14152	Sequence 14152, A
C 29	104	4.5	44120	3	US-09-949-016-14153	Sequence 14153, A
C 30	104	4.5	44120	3	US-09-949-016-14154	Sequence 14154, A
C 31	104	4.5	44120	3	US-09-949-016-14155	Sequence 14155, A
C 32	104	4.5	44120	3	US-09-949-016-14156	Sequence 14156, A
C 33	103.5	4.5	2670	3	US-09-902-540-517	Sequence 517, App
C 34	103	4.5	30001	2	US-08-125-468-1	Sequence 1, Appl
C 35	103	4.5	30001	2	US-08-474-933-1	Sequence 1, Appl
C 36	102.5	4.5	4695	3	US-09-902-540-8855	Sequence 8855, App
C 37	102.5	4.5	8658	3	US-09-902-540-943	Sequence 943, App
C 38	101.5	4.4	3059	3	US-09-902-540-540	Sequence 540, App
C 39	101.5	4.4	53615	3	US-09-214-808-1	Sequence 1, Appl
C 40	100.5	4.4	234884	3	US-09-949-016-16420	Sequence 16420, A
C 41	100	4.4	4041	3	US-09-949-016-1305	Sequence 1305, App
C 42	100	4.4	4041	3	US-09-949-016-1306	Sequence 1306, App
C 43	100	4.4	4416	3	US-09-919-039-116	Sequence 116, App
C 44	100	4.4	8930	3	US-09-077-098A-1	Sequence 1, Appl
C 45	100	4.4	8930	3	US-10-192-584-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-843-905A-7
Sequence 7, Application US/09843905A

Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843, 905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200, 198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patent version 3.1
SEQ ID NO 7
LENGTH: 1263
TYPE: DNA
ORGANISM: Homo sapiens
US-09-843-905A-7

Alignment Scores:

Pred. No.: 1,38e-238
Score: 2283.00
Percent Similarity: 99.8%
Best Local Similarity: 99.8%
Query Match: 99.7%
DB: 3
Gaps: 0

US-10-041-030-4 (1-420) x US-09-843-905A-7 (1-1263)

QY 1 MetPhaserProCjgIngluGluHhScCyAlaPrAsnLyGluLpVAllyrTyGly 20

```

Db      1 ATGTTTTCCCTGCGCAGAGAAACATGCGCCCAATAAGAGCCAGTGAATAACGGG 60
Qy      21 GluLeuValValLeuGlyTyrAenGlyAlaLeuProAenGlyAspArgGlyAArgGly 40
Db      61 GACCTGGTGGTCTCGGGTACAAATGGTGTCTTACCCAAATGAGATGAGAGACGAGGAAA 120
Qy      41 SerArgPheAlaLeuTyrIlyAspArgProLyAspAlaAenGlyValIlySProSerThrValHis 60
Db      121 AGTAGATTTGCTCTTACAAAGCGGCCCCAAGCAATGATGTCTCAAAACCCAGCACGCTCAT 180
Qy      61 ValIleSerThrProGlnAlaSerLyAlaIleSerCysAllyGlyIleHisSerIleSer 80
Db      181 GTGATATCCACGCGCCCGGAGCATCCAGGCTATCAGCTGCAGAAAGTCAACAGATATATCC 240
Qy      81 TyrThrLeuSerArgAenGlnThrValValIleGlyTyrThrHisAspLyAspThrAsp 100
Db      241 TACACCTTGTCAGAGAAATCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy      101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
Db      301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGCATCTCGTGTCAACAGACAGATT 360
Qy      121 SerGlySerGlnAenThrAspGlyAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361 TCTGCGCAGCAGAAACAGGAGCAGAGCCAGATCACAGAGACCAATATCCAGGTTCCGCC 420
Qy      141 CysArgIleValCysAspArgAenGlnProTyrThrIleArgIlePheAlaIleGlyPhe 160
Db      421 TGCAGAGTCGTGTGCGACAGAAATGAACCTTACACACAGAGAAATTTGCGCGCGATTT 480
Qy      161 AspSerSerLyAsenIlePheLeuGlyGlyIlyAlaIleAlaIleTyrIlyAspProAspGly 180
Db      481 GACTCTTCCAAAACATATTTCTTGAGTAAAGCAGCAAAAGTGAAGAAACCCCGACGCG 540
Qy      181 HisMetAspGlyLeuThrThrAenGlyValLeuValIleHisProArgIlyGlyPheThr 200
Db      541 CACATGATGGGCTCCTACTAATGAGCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy      201 GlnGlnSerGlnProGlyValIleTyrArgGlnIleSerValCysGlyAspValIleTyrThrLeu 220
Db      601 GAGGAGTCCCAAGCCCGGGGTGTGCGGAGATCTCTGTGTGTGTGTGTGTGTGTGTGTGT 660
Qy      221 ArgGlnThrArgSerAlaGlnGlnArgGlyIlyLeuValGlnSerGlnThrAspValIleu 240
Db      661 CGAGAAACAGGTGCGGCCAGCAACAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Qy      241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrArgThrAlaAspGly 260
Db      721 CAGGACGGCTCCCTCATTTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy      261 LeuPheHisThrProThrGlnIlyHisIleGlnAlaLeuArgGlnGlnIleAsnAlaIle 280
Db      781 CTTTTCATCTCCAACTCAGAACACATAGAACCTCCGCGAGGAGATTAAGCCGCC 840
Qy      281 ArgProGlnCysProValIlyLeuAsnThrLeuAlaPheProSerIleAsnArgLyGly 300
Db      841 CGGCTCAGTGTCTGTGGGCTCAACACCTGTGCTTCCCAACATCAACAGAAAGAG 900
Qy      301 ValValGlnGlyIlyGlnProTyrAlaTyrLeuSerCysGlyHisIleGlyTyrHis 320
Db      901 GTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Qy      321 AsnTyrGlyHisArgSerAspThrGlnAlaAenGlnArgGlyCysPrometCysArgThr 340
Db      961 AACCTGGGCGCATTCGAGGTGACACGAGAGCCCAACAGAGGAGGTCTCCCATGTGACAGACT 1020
Qy      341 ValGlyProTyrValProLeuTyrLeuGlyCysGlnAlaGlyPheTyrValAspAlaGly 360
Db      1021 GTGGGCGCCCTATGTGCTCTGTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Qy      361 ProProThrHisAlaPheThrProCysGlyHisIleValCysSerGlyIlySeralAlaIlyTyr 380
Db      1081 CCGGCAACTATGTCTTCACTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140

```

```

Qy      381 TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
Db      1141 TGGTCTCAGATCCCGTGTGCTCATGGAACCTATGATTTTCAAGCTCTTGCCCTTCTGT 1200
Qy      401 AlaThrGlnLeuValGlyGlnGlnAenCysIleIlyLeuIlePheGlnGlyProIleAsp 420
Db      1201 GTTACACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260

RESULT 2
US-09-843-905A-5
; Sequence 5, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843, 905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-843-905A-5

Alignment Scores:
Pred. No.: 7,05e-227 Length: 1260
Score: 2175,50 Matches: 399
Percent Similarity: 97.4% Conservative: 10
Best Local Similarity: 95.0% Mismatches: 10
Query Match: 95.0% Indels: 1
DB: Gaps: 1

US-10-041-030-4 (1-420) x US-09-843-905A-5 (1-1260)
Qy      1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAsnLyGlnProValIlyStryGly 20
Db      1 ATGTTTTCCCGGCGCAGAGAAACCCAGCGCCCAACAGAGCGGTGAATAATCGGG 60
Qy      21 GluLeuValValLeuGlyTyrAenGlyAlaLeuProAenGlyAspArgGlyAArgGly 40
Db      61 GACCTGGTGGTCTCGGGTACAAATGGTGTCTTACCCAAATGAGATGAGAGACGAGGAAA 120
Qy      41 SerArgPheAlaLeuTyrIlyAspArgProLyAspAlaAenGlyValIlySProSerThrValHis 60
Db      121 AGTAGATTTGCTCTTACAAAGCGGCCCCAAGCAATGATGTCTCAAAACCCAGCACGCTCAT 180
Qy      61 ValIleSerThrProGlnAlaSerLyAlaIleSerCysAllyGlyIleHisSerIleSer 80
Db      181 GTGATATCCACGCGCCCGGAGCATCCAGGCTATCAGCTGCAGAAAGTCAACAGATATATCC 240
Qy      81 TyrThrLeuSerArgAenGlnThrValValIleGlyTyrThrHisAspLyAspThrAsp 100
Db      241 TACACCTTGTCAGAGAAATCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy      101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
Db      301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGCATCTCGTGTCAACAGACAGATT 360
Qy      121 SerGlySerGlnAenThrAspGlyAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361 TCGGCGCGTCAAGAGAAAGAT--GCCAGATCACACAGAGCAACATCTCTAGTTGCGCA 417
Qy      141 CysArgIleValCysAspArgAenGlnProTyrThrIleArgIlePheAlaIleGlyPhe 160
Db      418 TGCAGAGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 477
Qy      161 AspSerSerLyAsenIlePheLeuGlyGlyIlyAlaIleAlaIleTyrIlyAspProAspGly 180

```



```

|||||
Db 478 GATTCTTCCAAAATATCTTCTTGAGAGAAAGCAAAATGGAAGAAACCTCGATGGA 537
|||
Qy 181 HiMeAspGlyLeuThrThraSnglyValLeuValMeHiSPRoArglyGlyPheThr 200
|||
Db 538 CACATGATGATGACTACTACCAATGGTGCTCTAGTGTGACCGCCGAGAGAGGCTTCACC 597
|||
Qy 201 GluGluSerGlnProGlyValTTPArgGluIleSerValCysGlyAspValTyrThrLeu 220
|||
Db 598 GAGGAATCCAGGCTGAGTGTGAGAGAAATCTCTGTGTGGGAAATGTACACCTTG 657
|||
Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
|||
Db 658 CGAGAGACCAAGTGTGGCCGAGAGAGAGGGAAGCTGTGTGAAGTGAACCAACGTCTCG 717
|||
Qy 241 GluAspGlySerLeuIleAspLeuCysGlyValAThrLeuLeuTPrArgThrAlaAspGly 260
|||
Db 718 CAAAGCGGCTCCCTCATGTGACCTGTGTGGGCACTCTCTCTGAGAAACCGCAGATGGC 777
|||
Qy 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
|||
Db 778 CTTTTCAGCGCTCTACTGACAAAGACATAGAAAGCCCTCCGCGAGAGATCAATGACGCC 837
|||
Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
|||
Db 838 CGACCCCAAGTCCCGGCGGCGCTTAACACCTTGCCCTCCCGACATCAACCGAGAGAA 897
|||
Qy 301 ValValGluGluLysGlnProTPrAlaTyrLeuSerCysGlyYhiSvalHisGlyTyrHis 320
|||
Db 898 GTGTGTGAAGAGAGAGAGCCCTGTGGGCAATCCTGAGCTGCGGCAATGTGACCGGCTAC 957
|||
Qy 321 AsnTPrGlyYhiArgSerAspThrGluAlaAsnGluArgLysProMetCysArgThr 340
|||
Db 958 AGCTGGGGCCATGGAGCGAGCGGAGAACCAAGAGAGAGATGTCCCATGTGCAGAGACT 1017
|||
Qy 341 ValGlyProTyrValProLeuTPrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
|||
Db 1018 GTGGGCCCCCTACCTCTCTCTGTGGCTGTGAGGACAGATTTTATGTGATGCGGGA 1077
|||
Qy 361 ProProThrIleAlaPheThrProCysGlyYhiSvalCysSerGluYsSerAlaYsTyr 380
|||
Db 1078 CCCCACCACTCAGCTTTCACCCCCTGGGCGACGCTGTTCAGAAAGTCTGCGCAAGTAC 1137
|||
Qy 381 TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
|||
Db 1138 TGTGTCCAGATCCCACTGCCCAAGCAAGCAAGCGCTTTCATGCGGCTGTCTGTGC 1197
|||
Qy 401 AlaThrGlnLeuValGlyGluGlnAsnCysAlleuValLeuIlePheGlnGlyProIleAsp 420
|||
Db 1198 GCCACGACGCTGGTGTGTGAAACAGAACTGCATCAAAATTGATTTCCAAAGTCCAGTGGAC 1257
|||
RESULT 3
US-09-843-905A-3 ; Sequence 3, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843, 905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-843-905A-3
Alignment Scores:

```

```

Pred. No.: 1,02e-198 Length: 1257
Score: 1917.00 Matches: 342
Percent Similarity: 90.5% Conservative: 38
Best Local Similarity: 81.4% Mismatches: 38
Query Match: 83.7% Gaps: 2
DB: 3 Indels: 1

US-10-041-030-4 (1-420) x US-09-843-905A-3 (1-1257)

Qy 1 MecPheSerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
|||
Db 1 ATGTTTCTCTCTGATCAAGAAATCAT-----CCATCTAAACACACAGATAAATATGTGT 54
|||
Qy 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyValArgLys 40
|||
Db 55 GAACATATGTCTTAGAATATATGATGATCTCTCCAAACGGTATATGAGAGAGAGAGAAA 114
|||
Qy 41 SerArgPheAlaLeuTyrLysArgProLysAlaAlaIleSerCysLysGlyGlnHisSerIleSer 60
|||
Db 115 AGTAGTTCCTTGTGTTTAAAGACCTTAGGCAAAATGGGTGAAGCCGACCTGTGCAT 174
|||
Qy 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
|||
Db 175 ATGCTTGTACTCTCTCAGGCTGCAAGGCAATAGCAACAAAGACGATACATATCA 234
|||
Qy 81 TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAspLysAspThrAsp 100
|||
Db 235 TATACTTTATCTCGGGCCGACGACTGTGTGTGAATATATCATATACACCAACCCGAT 294
|||
Qy 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
|||
Db 295 ATGTTCAGATTGCGCGGTGCACTGAAAGCCCATTTGATTTGTAGTACGACGGTT 354
|||
Qy 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
|||
Db 355 CCGGAAGTCAAAAGTATCTGTATACACAGTCAAGTCAAAAGCATATATCAAGATTGGCC 414
|||
Qy 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
|||
Db 415 TCGAAGATCATATGTGAACGGAATCCTCTCTTACAGCAGGATTTATGTGCGAGGGTTT 474
|||
Qy 161 AspSerSerLysAsnIlePheLeuGlyLysValAlaAlaLysTPrLysAsnProAspGly 180
|||
Db 475 GACTCATCAAAAACATCTTTCTTGAGAGAGAGGCTGCCAAATGAGAAAGACATCAAGTGA 534
|||
Qy 181 HiMeAspGlyLeuThrThraSnglyValLeuValMeHiSPRoArglyGlyPheThr 200
|||
Db 535 CACATGATGATGCTTGAACCACTAATGATGTTCTTGTATGATGATCAAGCAATGAGTTTACA 594
|||
Qy 201 GluGluSerGlnProGlyValTTPArgGluIleSerValCysGlyAspValTyrThrLeu 220
|||
Db 595 GAAGACTCCAGCGCTGGAATATGAGAAATATCGGTGTGTGGAAATGATTAATTAAGCTTA 654
|||
Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
|||
Db 655 CGTGAACCAAGTGTGCTGACAGAGAGAGAAATGTGAAATTTAAACCAATTCAGTTA 714
|||
Qy 241 GluAspGlySerLeuIleAspLeuCysGlyValAThrLeuLeuTPrArgThrAlaAspGly 260
|||
Db 715 CAAAGTGTGCTGTATTAATTAACCTGTGTGAGCAACTGTTTAAGGCTACTGCAAGAGGC 774
|||
Qy 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
|||
Db 775 CTTTCCACACCTCTCAGCGGAGCAATTTAAGAGCTTTAAGACAGAAATCAATGACGCA 834
|||
Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
|||
Db 835 CGACCTCAGTCCCTTAGAGGTTCAACACATGCAATTCCTAGTATGAAAGAGAGAAAGAC 894
|||
Qy 301 ValValGluGluLysGlnProTPrAlaTyrLeuSerCysGlyYhiSvalHisGlyTyrHis 320
|||
Db 895 GTTGTAGTGAAGAAACACACATGATATCTTAATCTGCGGCAATGATACATGCGTATCAT 954
|||

```

```

QY 321 AaNTpGlyHIsaRgSerAepThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
DB 955 AACTGGGAACAAGAAAGACGTATGCGAAGATGTGATGTCTTATGTGTAGGTCT 1014
QY 341 ValGlyProTyValProLeuTrpLeuGlyCysGluAlaGlyPheTyValAspAlaGly 360
DB 1015 GTTGGTCCCTATGTCTCTGTGGCTGGATGTGAAGCTGATTTATGTGACGCCGGC 1074
QY 361 ProProThHIsAlaPheThrProCysGlyHIsValCysSerGluYsSerAlaTy 380
DB 1075 CCTCCAAACCATGCTTACCGCTGTGGCATGTGTCTTCAAGAAAGACACTGCTAT 1134
QY 381 TrpSerGlnIleProLeuProHIsGlyThrHIsAlaPheHIsAlaAlaCysProPheCys 400
DB 1135 TGGTCCCAAGATCCCATCTCTCATGTGATCATATCTTTCATGACGCTGTCTTTGT 1194
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
DB 1195 GCACATCAGTTGGCTGGTGAACAAGCTACATCAGACTTATTTTTCAGAGACCTTAGAC 1254

RESULT 4
US-09-843-905A-1
; Sequence 1, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843, 905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-843-905A-1

Alignment Scores:
Pred. No.: 3,57e-198 Length: 1257
Score: 1912.00 Matches: 341
Percent Similarity: 90.5% Conservative: 39
Beet Local Similarity: 81.2% Mismatches: 38
Query Match: 83.5% Indels: 2
DB: 3 Gaps: 1

US-10-041-030-4 (1-420) x US-09-843-905A-1 (1-1257)
QY 1 MetPheSerProGlyGlnGluGlnHIsCysAlaProAsnIlySerProValIleTyGly 20
DB 1 ATGTTTCTCTCTATCAAGAAATCAT-----CCTTCAAGGCCCAAGTAAATATAGGC 54
QY 21 GluLeuValIleuGlyTyraGnGlyAlaLeuProAsnGlyAspArgGlyArgGly 40
DB 55 GAATCTATTGTCTTGAATTAATGATCTCTCCAAACGGTATGAGAGAGAGAGAA 114
QY 41 SerArgPheAlaLeuTylyArgProIlyAsnGlyValIlyAspProSerThrValHIs 60
DB 115 AGTAGGTTTCTTGTAAAGCCTAAGCAAAATGGGTGAAGCTTACACCGTGCAC 174
QY 61 ValIleSerThrProGlnAlaSerIlyAlaIleSerCysAllyGlyGlnHIsSerIleSer 80
DB 175 ATTCATGTATCTCTCAGCTGCGCAAGCAATTAAGCAACAGACCATAGCATATCA 234
QY 81 TyThrIleuSerArgAsnGlnThrValValIleGlyTyThrHIsAspIlyAspThrAsp 100
DB 235 TATACCTTATCTGAGCCCAAGACGCTGTGTGAATATCTCTGACACCAACACAGAT 294
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
DB 1015 GTTGGTCCCTATGTCTCTGTGGCTGGATGTGAAGCTGATTTATGTGACGCCGGC 1074

```

```

DB 295 ATGTTTCAGATGGTGGTCAACTGAAGCTCATTTGATTTGTATGAACGACCGTT 354
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 355 CCGGAAGTCAAGGTATTCGACACGACGTCAATGACATTAACAGATTTGCGC 414
QY 141 CysArgIleValCysAspArgAsnGlyProTyThrAlaArgIlePheAlaAlaGlyPhe 160
DB 415 TGTAGATCATATGTGAGCGGACGTCCCTTTACAGCTCGGATTTATCTGTGACGGGTTT 474
QY 161 AspSerSerIleAsnIlePheLeuGlyGlyValAlaAlaIleTyTrpIlyAsnProAspGly 180
DB 475 GATTCATCAAAAACATCTTCTTGGGAGAAAGGCTCCAAAGTGAAGACATCTGATGGG 534
QY 181 HIsMetAspGlyLeuThrThrAsnGlyValIleValIleMetHIsAspArgGlyGlyPheThr 200
DB 535 CAGATGATGGCTTGAACCATTAATGAGATTTCTTGTATGATCAATCAGTAAATGGGTTCA 594
QY 201 GluGluSerGlnProGlyValTrpArgGluIleSerValCysGlyAspValTyThrLeu 220
DB 595 GAAGACTCCAAACCTGGAATATGAGAGAAATATCATGATGTGGGAATGCTTCACTG 654
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyIlyLeuValGluSerGluThrAsnValLeu 240
DB 655 CCGTGAACCAAGATCACTCAGACAGAGAGAAAGATGTGAAATTGAACCAATTCAGCTA 714
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuThrTrpArgThrAlaAspGly 260
DB 715 CAAGATGGCTCTTAATGACCTTGTGTGGTCACTGTGTGGTCTGTGCTGCTGCAAGAGC 774
QY 261 LeuPheHIsThrProThrGlnIlyHIsIleGlyAlaLeuArgGlnIlyIleAsnAlaAla 280
DB 775 CTTTCCCATATCTCTATCTGTGAAGACATTAAGAGCTTTAAGACAGAGATCATGACGCT 834
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
DB 835 CGGCCCGACGTGGCTGTAGGGTTCAACACATGACCTTCCCGTGTGAAGAGAGAT 894
QY 301 ValValGluGluIlySerProTrpAlaTyThrLeuSerCysGlyHIsValHIsGlyTyHIs 320
DB 895 GTTGTAGATGAAGAACCAATGAGTATATCTTAACCTGGCGCATGTCCATGGGTATCAT 954
QY 321 AaNTpGlyHIsaRgSerAepThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
DB 955 AACTGGGAACAAGAAAGACGTATGCGAAGATGTGATGTCTTATGTGTAGGTCT 1014
QY 341 ValGlyProTyValProLeuTrpLeuGlyCysGluAlaGlyPheTyValAspAlaGly 360
DB 1015 GTTGGTCCCTATGTCTCTGTGGCTGGATGTGAAGCTGATTTATGTGACGCCGGC 1074
QY 361 ProProThHIsAlaPheThrProCysGlyHIsValCysSerGluYsSerAlaTy 380
DB 1075 CCTCCAAACCATGCTTACCGCTGTGGCATGTGTCTTCAAGAAAGACACTGCTAT 1134
QY 381 TrpSerGlnIleProLeuProHIsGlyThrHIsAlaPheHIsAlaAlaCysProPheCys 400
DB 1135 TGGTCCCAAGATCCCATCTCTCATGTGATCATATCTTTCATGACGCTGTCTTTGT 1194
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
DB 1195 GCACATCAGTTGGCTGGTGAACAAGCTACATCAGACTTATTTTTCAGAGACCTTAGAC 1254

RESULT 5
US-09-843-905A-11
; Sequence 11, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843, 905A
; CURRENT FILING DATE: 2001-04-27

```

PRIOR APPLICATION NUMBER: US 60/200,198
 PRIOR FILING DATE: 2000-04-28
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 11
 LENGTH: 1338
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (513)..(513)
 OTHER INFORMATION: unsure
 US-09-843-905A-11

Alignment Scores:
 Pred. No.: 1,02e-168 Length: 1338
 Score: 1642.00 Matches: 299
 Percent Similarity: 81.4% Conservative: 42
 Best Local Similarity: 71.4% Mismatches: 74
 Query Match: 71.7% Indels: 4
 DB: 3 Gaps: 3

US-10-041-030-4 (1-420) x US-09-843-905A-11 (1-1338)

```

QY 3 SerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGlyLysLeu 22
DB 88 TCTCCCGGT-----GAAAGATGCCGACGCCGAGGAGGCCCAATCAAGTATGTGAATC 141
QY 23 ValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgATGlySerArg 42
DB 142 ATCGTCTGGGCTTCACATGTGTCTGTGCAAGTGGGACAAAGGCGCGCCGCAAGCGCG 201
QY 43 PheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHisValIle 62
DB 202 CTGGCAGCTGAGCGCGCTGCGACGCGCAAGCGGCTGAAGCAGACGTCATGACACATC 261
QY 63 SerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerLysSerTyrThr 82
DB 262 TCCAGCGCGCTGCTCTCCAAAGGACCTGATACCGTGTGTGACGACGACATCTGATACA 321
QY 83 LeuSerArgAsnGlnThrValAlaValGluTyrThrHisAspLysAspThrAspMetPhe 102
DB 322 CTGTCCCGGAGCCACTCGCTCATAGTGATATACATATACATATACGACACACATCTTC 361
QY 103 GlnValGlyArgSerThrGlnSerProLysAspPheValAlaThrAspThrLysSerGly 122
DB 382 CAGATTGGCGGCTCCACAGAGAACATGATGACTTCGTGTAAACAGACAGTCCCGTGA 441
QY 123 SerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrLysSerArgPheAlaCysArg 142
DB 442 GGA---GGGGCTGCCGAGGGCCCTTCTGCCAGAGCACATCTCCCGCTATGCTGCCG 498
QY 143 IleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPheAspSer 162
DB 499 ATCTCTGTACACCGGCGCCACCTTATCTGCGGACATCTATGCGCTGCTGCTGATGCC 558
QY 163 SerLysAsnIlePheLeuGlyGlyLysAlaAlaLysTyrLysAsnProAspGlyHisMet 182
DB 559 TTAGAGCAACATCTTCTTGGAGGCGAGCGGCGCAAAATGGGAGACCCAGATGGCGTGA 618
QY 183 AspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyIleThrGlnGlu 202
DB 619 GATGAGCTGACCAACATGAGATGCTGTGATGACCCGCGAGCGGCTTCTCCGAGAGAC 678
QY 203 SerGlnProGlyValTyrArgGluLysSerValCysGlyAspValTyrThrLeuArgGlu 222
DB 679 TCGAGCCCGGCTGTCTGCGGAGATCTCGCTCTGTGGAATGTATCAATTCGCGGAC 738
QY 223 ThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeuGlnAsp 242
DB 739 AGCGCGTCAAGCCAGAGCGGCGGAGAGCTGTGAAAAACAGATCCAAAGTCTGCGAGAC 798
QY 243 GlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrPargThrAlaAspGlyLeuPhe 262
  
```

```

DB 799 GGCTCTCATGACCTGTGTGGGGGACACATGCTGTGGGACACACCGGCGGGCTGTG 858
QY 263 HisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAlaArgPro 282
DB 859 CGGGCTCCCACTGTAAGCAACTGAGGGCCGACGGCGCAGAGAGCAAAATGACGCGCC 918
QY 283 GlnCysProValGlyLeuAsnThrLeuAlaPheProSerLysAsnArgLysGluValAla 302
DB 919 CAGTGGCCCGTGGGCTTCAGCACTGTGGCTTCTCCAGGCCAGCCCGCGGACAGCG 978
QY 303 GlnGluLys---GlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsn 321
DB 979 CCGGACAAACAGAGGCGCTGTGCTACGTCGCGGAGACGTCATGCTACACAGCGC 1038
QY 322 TrrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThrVal 341
DB 1039 TGGGGCTGGCGGCGGAGCGGCGCCCGCAGAGCGCGAAATGTCTCTGCGCCCTGTG 1098
QY 342 GlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGlyPro 361
DB 1099 GGGCTTATGTGCTCTATGTGCTTGGCCAGAGGCGCGCTTGTGCTGACCTTGAGCG 1158
QY 362 ProThrHisAlaPheThrProCysGlyHisValCysSerGlyLysSerAlaLysTyrTrp 381
DB 1159 CTTAGCCATGCTTTGACACTTGGCGGACGCTGCTGTGAAAGACTGCGGCTACTG 1218
QY 382 SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCysAla 401
DB 1219 GCCCAAGACCACTGGCCCAAGGACCAATGCTTTCATAGCGCGCTGCGCCCTTTGCGGG 1278
QY 402 ThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
DB 1279 GCCTGCTTACCGGCGAGATGCGCTGCGCTCATTTTCCAGGCGCGCTGAT 1335
  
```

RESULT 6
 US-10-104-047-1958
 Sequence 1958, Application US/10104047
 Patent No. 6943241
 GENERAL INFORMATION:
 APPLICANT: HELIX RESEARCH INSTITUTE
 TITLE OF INVENTION: No. 6943241el full length cdna
 FILE REFERENCE: H1-A0105
 CURRENT APPLICATION NUMBER: US/10/104,047
 PRIOR FILING DATE: 2002-03-25
 PRIOR APPLICATION NUMBER:
 NUMBER OF SEQ ID NOS: 4096
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1958
 LENGTH: 2508
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-104-047-1958

Alignment Scores:
 Pred. No.: 6.73e-151 Length: 2508
 Score: 1482.50 Matches: 275
 Percent Similarity: 74.2% Conservative: 36
 Best Local Similarity: 65.6% Mismatches: 65
 Query Match: 64.7% Indels: 43
 DB: 3 Gaps: 4

US-10-041-030-4 (1-420) x US-10-104-047-1958 (1-2508)

```

QY 3 SerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGlyLysLeu 22
DB 208 TCTCCCGGT-----GAAAGATGCCGACGCCGAGGAGGCCCAATCAAGTATGTGAATC 261
QY 23 ValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgATGlySerArg 42
DB 261 ----- 261
  
```

QY 43 PheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHisValIle 62
Db 262 -----ATC 264
QY 63 SerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSerTyrThr 82
Db 265 TCACACGCGCTGCTCTCCAGAGCAGTGAACCGTGTGATGACAGACATCTGATACA 324
QY 83 LeuSerArgAsnGlnThrValValGluTyrThrHisAspLysAspThrAspMetPhe 102
Db 325 CTCTCCCGAGCAGCTGCTGATAGTGAATACATATATGACGACAGACATGTTTC 384
QY 103 GlnValGlyArgSerThrGlnSerProIleAspPheValValThrAspThrIleSerGly 122
Db 385 CAGATTGCGCGCTCCACAGAGAACATGATGATCTCTGTGAACAGACAGTCCCGGA 444
QY 123 SerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArg 142
Db 445 GGA---GGGCTGCGGAGGCGCTTTCGCCAGAGCACATCTCCGCTATGCTGCGCG 501
QY 143 IleValCysAspArgAsnGlnProTyrThrAlaArgIlePheAlaAlaGlyPheAspSer 162
Db 502 ATCTCTGTACCGCGCCAGCCTTATATGCGCGCATCTATGCGCTGCTGATGCC 561
QY 163 SerLysAsnIlePheLeuGlyGluLysAlaAlaLysTyrLysAsnProAspGlyHisMet 182
Db 562 TCTAGCAACATCTTCTTGTGAGACCGAGCGCCAAATGCGGAGACCCAGATGCGCTGATG 621
QY 183 AspGlyLeuThrThrAsnGlyValLysValMetHisAspArgGlyGlyPheThrGlu 202
Db 622 GATGAGTACGACCAACATGAGTCTGTGATGACCGGAGCGGCGCTTCTCCAGAGAC 681
QY 203 SerGlnProGlyValTyrPargGlnIleSerValCysGlyAspValTyrThrLeuArgLys 222
Db 682 TCAGCCCGCGGTGTGCGGAGATCTCGTGTGAGATGATGATGATGATGATGATGATGATG 741
QY 223 ThrArgSerAlaGlnGlnArgGlyLysLeuValGlnSerGlnThrAsnValLysGlnAsp 242
Db 742 AGCGCTCAAGCCAGCGCGGCGGAGACTGTGTGAACAAAGTCCAACTGCTGACAGAC 801
QY 243 GlySerLeuIleAspLeuGlyValIleThrLeuLeuTyrThrAlaAspGlyLeuPhe 262
Db 802 GGCTCTCATGACCTGTGTGCGGACACTGTGTGCGGACACCGGCGGCGCTGCTG 861
QY 263 HisThrProThrGlnLysHisIleGlnAlaLeuArgGlnGlnIleAsnAlaAlaArgPro 282
Db 862 CGGCGCTCCACACTGAAGCAACTGAGAGCCGAGCGAGGAGGCAATGCAAGCGCGGCC 921
QY 283 GlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValVal 302
Db 922 CAGTGGCGGCTGCGCTGACGACTGTGCTTCCCGAGCCGCGGCTGCGCGGAGCG 981
QY 303 GlnGluLys---GlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsn 321
Db 982 CCGAGCAAAACAGAGCGCTGCGTCTACGTCCGCGCGGAGCAGTCCAGTGAACAGCGC 1041
QY 322 TyrGlyHisArgSerAspThrGlnAlaAsnGluArgLysProMetCysArgThrVal 341
Db 1042 TGGGGCTGCGCGGAGCGGCGGCGCGCGAGGCGCAATGCTCTCTGCGCGCTGCTG 1101
QY 342 GlnProTyrValProLeuTyrPheGlyCysGlnAlaGlyPheTyrValAlaAspAlaGlyPro 361
Db 1102 GGGGCTATGCTCTATGCTGTGCGAGAGGCGCGCTGCTGCGAGCGCTGCGCGCG 1161
QY 362 ProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyrTyr 381
Db 1162 CCAAGCATGCTTGTGACCTGCGGAGCAGTCTGCTGAGAGAGATGCGCGCTATCTG 1221
QY 382 SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCysAla 401
Db 1222 GCCCAGACACACAGCGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 1281
QY 402 ThrGlnLeuValGlyGlnGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420

Db 1282 GCGTGGCTTACCGGAGCATGCGTGTGCGGCTCATTTTCCAGGCGCGTGTGAT 1338
RESULT 7
US-08-136-277-1
Sequence 1, Application US/08136277
Patent No. 5644045
GENERAL INFORMATION:
APPLICANT: MANDEL, Jean-Louis
APPLICANT: AUBOURG, Patrick
APPLICANT: MOSSER, Jean
APPLICANT: SARDE, Claude
TITLE OF INVENTION: X-LINKED ADRENOLRUKODYSTROPHY GENE AND
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Young & Thompson
STREET: 745 South 23rd Street
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,277
FILING DATE: 15-OCT-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: B2272
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBOW
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2750 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 387..2624
US-08-136-277-1
Alignment Scores:
Pred. No.: 0.0186 Length: 2750
Score: 120.50 Matches: 104
Percent Similarity: 33.9% Conservative: 63
Best Local Similarity: 21.1% Mismatches: 155
Query Match: 5.3% Indels: 171
Gaps: 26
US-10-041-030-4 (1-420) x US-08-136-277-1 (1-2750)
QY 2 PheSerProGlyGlnGlnHisCysAlaProAsnLysGluProValLysTyrGlyGlu 21
Db 1167 TTCTCGGCC-----AAGTTGCGGAG 1187
QY 22 LeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgLysSer 41
Db 1188 CTGCTG-----CGACAGAGCGCGCGCGGAGCGG 1217
QY 42 ArgPheAla---LeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 1218 GAGCTGGCTATACATGACATGCTGCTGTGTGCGGCAAC----- 1253

Qy 61 ValIleSerThrProGlnAlaSerIleValIleSerCysIleGlyGlnHisIleSerIleSer 80
 Db 1254 -----TCGAGAGAGATGCGCTTCTATGAGGGGCGCATGAGTGGAG 1292
 Qy 81 TyrThrIleu-----SerArgAsnGlnThrValVal 91
 Db 1293 CTGGCCCTGTACAGCCCTCTACAGAGACCTGGCTCGCAGATCAACATCATCTCTCTG 1352
 Qy 92 Gln-----TyrThrHisAspIleAspThrAspMetPheGlnValGlyArgSerThr 108
 Db 1353 GAACGCGCTGTGTATGTATGCTGAGCAGAGTCTCTCAAGATGATGTGTGAGAGCGCTCG 1412
 Qy 109 GluSerProIleAspPheValIleThrAspThrIleSerGlySerGlnAsnThrAspGlu 128
 Db 1413 GCGCTGTCTATGCGGTGTCTCCCATCATCATCTGCTCACTGCTCTCAAGTCTGATGCA 1472
 Qy 129 AlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAspArgAsn 148
 Db 1473 GAGGCGGTGAAG 1532
 Qy 149 GluProIleThr-----AlaArgIlePheAlaIleAlaIlePheAspSerIleValAsnIle 166
 Db 1533 GAAGCTCTCACTATTTGCCCGCACTCTGACAGCGCTGCAATGCGATTCAGGAGT 1592
 Qy 167 -----PheLeuGlyGlyValAlaIleValIleValIleValIleValIleValIleVal 183
 Db 1593 ATGTCGTGTACAA-GAAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1651
 Qy 184 -----GlyLeu----- 185
 Db 1652 GTTCAGAGTATTTGAAGATGTTGAGCGCTGTCACTTCAAGAGCCGAGAGAGAGAGAGAG 1711
 Qy 186 ThrThrAsnGlyValIleu-----ValMetHisProAspArgIleGlyPheThrGlu 201
 Db 1712 CGCTCAGCGCGGCTGTGAGCAGCATAGGCGGCTGTGAGTGTGAGAGAGAGAGAGAGAG 1771
 Qy 202 GluSerGlnProGlyValIleThrArgGlnIleSerValCysGlyAspValIleThrLeuArg 221
 Db 1772 GATCCAGGCGCAGGT-----GTTGATGTGTGAACAGAGAGAGAGAGATCTCTCGA 1816
 Qy 222 Gln-ThrArgSerAlaGlnIleArgIleGlyValLeuValGluSerIleThrAsnVal--Le 240
 Db 1817 GAACATCCCATGTCAGCGCCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1876
 Qy 240 uGlnAspGly----- 246
 Db 1877 GAG 1936
 Qy 246 eaAspLeuCysGlyAlaThrIleuLeuThrArgThrAlaAspGlyLeuPheHisThrProth 266
 Db 1937 CCGGATCTGAGTGGG-----CTCTGAGCCCACTAGCGGTGTGTCTTCAAGAGAGAGAG 1990
 Qy 266 rgIleuHisIleGlnAlaLeuArgGlnIleValAsnAlaAlaArgProGlnCysProva 286
 Db 1991 ACCCCAGGAGATGTTCAATCCCGCAG-----AGGCCCTCAATGTCGT 2035
 Qy 286 IGIY-----LeuAsnThrIleuAlaPheProSerIleAsnArgIleGlyValValGluG 304
 Db 2036 GGGCTCCCTCGTGAACAGAGATCTACCG-----GACTCAGTGAGAGA 2080
 Qy 304 uIysGlnProIleAlaIle-----LeuSerCysGlyHisIleVal 316
 Db 2081 CATGCAAG 2140
 Qy 316 IHis-----GlyTyrHis-----AsnThrGlyHis--Ar 325
 Db 2141 GCACCAATCTGTGAG 2200
 Qy 325 gSerAspThrGlu-AlaAsnGlnArgIleCysProMetCysArgThrValGlyProTyr 345
 Db 2201 GTGCGGTGTGAG 2260
 Qy 345 alProleuThrPleuGlyCysGlnAlaGlyPheThrValAlaPheAlaGlyPro----- 361

Db 2261 CGCCTCTGTGAG 2299
 Qy 362 -----ProThr 364
 Db 2300 CTGGAAGAGAGAGATCTTCCAGCGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2359
 Qy 364 IAlaPheThrProCysGly-----HisValCys 375
 Db 2360 CCACCGGCGCTCTCTGTGAAATATCCACACACACTTGTCAAGTTGATGGAGAGAGAG 2419
 Qy 375 IuIysSerAlaIleValIleThrSerGlnIlePro 385
 Db 2420 CTGGAAGTGTGAG 2451

RESULT 8
 US-08-479-403-1
 Sequence 1, Application US/08479403
 Patent No. 5869039
 GENERAL INFORMATION:
 APPLICANT: MANDEL, Jean-Louis
 APPLICANT: AUBOURG, Patrick
 APPLICANT: MOSSE, Jean
 APPLICANT: SARDE, Claude
 TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
 TITLE OF INVENTION: CORRESPONDING PROTEIN
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Young & Thompson
 STREET: 745 South 23rd Street
 CITY: Arlington
 STATE: VA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/479,403
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: PATCH, Andrew J.
 REGISTRATION NUMBER: 32,925
 REFERENCE/DOCKET NUMBER: B2272D1V
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-521-2297
 TELEFAX: 703-685-0573
 TELEX: 248425 EMBON
 INFORMATION FOR SEO ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2750 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 387..2624
 US-08-479-403-1
 Alignment Scores:
 Pred. No.: 0.0186 Length: 2750
 Score: 120.50 Matches: 104
 Percent Similarity: 33.9% Conservative: 63
 Best Local Similarity: 21.1% Mismatches: 155
 Query Match: 5.3% Indels: 171
 DB: 2 Gaps: 26
 US-10-041-030-4 (1-420) x US-08-479-403-1 (1-2750)

```

Qy      2 PhaserProgluIngluInHsCysAlaProAsnLysgluProValLysTyrGlu 21
Db      1167 TTCTCGCC-----AGTTGGGGAG 1187
Qy      22 LeuValValLeuGlyTyrAsnGlyAlaLeuProAsnLysAspArgGlyAlaGlyLysSer 41
Db      1188 CTGGTG-----GCAGAGGAGCGCGCGGAGGG 1217
Qy      42 ArgPheAla---LeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db      1218 GAGTCGGCTACATGACATCGCTGTGTGGCCAAAC----- 1253
Qy      61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysgluInHsSerIleSer 80
Db      1254 -----TCGAGAGAGTGCCTTCTATGGGGGCGCATGAGGTGAG 1292
Qy      81 TyrThrLeu-----SerArgAsnGlnThrValVal 91
Db      1293 CTGGCCCTGTACAGCCCTCTACAGAGACCTGGCCCTGCGAGATCAACTATCTTCTG 1352
Qy      92 Glu-----TyrThrHisAspLysAspThrAspMetPheGlnValGlyArgSerThr 108
Db      1353 GAAAGCCCTGGTATGTATGTCTGGAGCAAGTTCTCTCAAGATATGTGTGGAGCGCTCG 1412
Qy      109 GluSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu 128
Db      1413 GGCCTGCTCANTGGCTGTCCCATCATCATCTGCTGCTCACTGAGAGTCAGATGCA 1472
Qy      129 AlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAspArgAsn 148
Db      1473 GAGCGCCGTGAAGAGCAGCGCTTGGAAGAGAGAGAGAGAGCTGTGAGCGAGCGCACA 1532
Qy      149 GluProTyrThr-----AlaArgIlePheAlaIleGlyPheAspSerSerLysAsnIle 166
Db      1533 GAAAGCCCTCACTATTGCCCGCAACTCTCTGACAGCGCTGAGTGCATTGAGCGGATC 1592
Qy      167 -----PheLysGlyLysAlaAlaLysTyrLysAsnProAspArgLysIleMetAsp 183
Db      1593 ATGTCGTCGACAA--GGAAGTGAAGAGCTGGCTGGCTACAGACCGCGGTGACAGAT 1651
Qy      184 -----GlyLeu----- 185
Db      1652 GTTCCAGATATTTGAAGATGTTCAAGCGCTGCACTTCAAGAGCGCCAGGAGTACAGGA 1711
Qy      186 ThrThrAsnGlyValLeu-----ValMetHisProAspArgGlyLysPheThrGlu 201
Db      1712 CGCTCAGCGCGGCTGGAGCATAGCGCGGTCTGGTCCGTGGAGAGCGCCCTGAA 1771
Qy      202 GluSerGlnProGlyValTTPArgGluIleSerValCysGlyAspValTyrThrLeuArg 221
Db      1772 GATCCGAGGCGAGGT-----GTTGATGTGGAAACAGGGGATCATCTCGCA 1816
Qy      222 Glu-ThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnVal---Le 240
Db      1817 GAACATGCCCATGTCACGCCCTCAGAGAGAGTGTGTGGCCAGCTCAACATCAGAGGT 1876
Qy      240 uGlnAspGly-----SerLeuIle 246
Db      1877 GAGAGAGGAGCATGATCTGTCATACAGGCCCAATGGCTGCGCGCAAGAGCTCCCTGT 1936
Qy      246 eAspLysCysGlyAlaThrLeuLeuTTPArgThrAlaAspGlyLeuPheHisThrProth 266
Db      1937 CCGGATCTGTGGTGG-----CTCTGACCCCACTACAGTGTGTCTTACAAAGCCCTC 1990
Qy      266 rGlnLysHisIleGlnAlaLeuArgGlnGluIleAsnAlaAlaArgProGlnCysProva 286
Db      1991 ACCCCAGGAGCATGTTTCAATCCGCGAG-----AGGCGCTACATGCTCTGT 2035
Qy      286 IglY-----LeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValValGluI 304
Db      2036 GGGCTCCCTGCTGACCAAGGATCTACCCG-----GACTCAGAGAGGA 2080
Qy      304 uLysGlnProTTPAlaLys-----LeuSerCysGlyHisIleVa 316

```

```

Db      2081 CATGCAAGAGAGGCTACTCGAGAGGACCTGAAAGCATCTGAGACGTCTGCACCT 2140
Qy      316 IHis-----GlyTyrHis-----AsnTyrGlyHis---Ar 325
Db      2141 GCACCAATCTTCACCGGAGGAGGAGGTGGAGGCTATGTGTGATGGAGAGACCTCT 2200
Qy      325 gSerAspThrGlu-AlaAsnGluArgGlyLysProMetCysArgThrValGlyProTyrV 345
Db      2201 GTGGGTGGGGAAGACAGAGAAATCGGCATGGCCCGCATCTTCTTACACAGGCCCAAGTA 2260
Qy      345 aLProLeuTyrLeuGlyCysGluAlaGlyPheTyrValAspAlaGlyPro----- 361
Db      2261 CGCCCTTCGATGAATGACACACGCG-----CCGTGAGCATCGA 2299
Qy      362 -----ProThr 364
Db      2300 CGTGAAAGGAGATCTTCCAGGCGGCAAGAGAGCGGCGCATTCCTGCTTCATCAC 2359
Qy      364 IAlaPheThrProCysGly-----HisValCys-----SerG 375
Db      2360 CCAACGGCCCTCCCTGTGAAATACACACACACTTCTTACAGTTGATGGAGGAGCGG 2419
Qy      375 LuLysSerAlaLysTyrTTPSerGlnIlePro 385
Db      2420 CTGAAGTTCGAGAGAGCTGAGCTCAGCTGCC 2451

RESULT 9
US-08-835-734-1
Sequence 1, Application US/08835734
Patent No. 6013769
GENERAL INFORMATION:
APPLICANT: MANDEL, Jean-Louis
APPLICANT: AUBOURG, Patrick
APPLICANT: MOSSER, Jean
APPLICANT: SARDE, Claude
TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
TITLE OF INVENTION: CORRESPONDING PROTEIN
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Young & Thompson
STREET: 745 South 23rd Street
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,734
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/479,403
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: B2272D1V
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SBO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2750 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

```


Query Match: 5.0% Indels: 129
DB: 3 Gaps: 21
US-10-041-030-4 (1-420) x US-09-103-840A-2 (1-4403765)
QY 4 ProGlyGlnGlu-----GluHieCyAlaProAnlybgluProValLysTyr 19
Db 1250835 CCTGGGCAACAAAGGCGCAGCATGACATGATCGCGGCGACCATCTGCC----- 1250785
QY 20 GlyGluLeuValValLeuGlyTyrAenGlyAlaLeuProAnGlyAaPaRgGlyAaRg 39
Db 1250784 GGTGGCGCGCTGCAATTCGATTACAC-----CGAGGTGGCGGT 1250746
QY 40 LysSerArgPheAlaLeuTyrLysArgProLysAlaAenGlyValLysPro----- 56
Db 1250745 -----GGCTATCTCGCGCGCGCGGCAATCCAGATCCCGCGCATCGGATGA 1250698
QY 57 -----SerThyValHisValLieserThPro 65
Db 1250697 CGTGTGATCGATGATGATGATGATGATAAGAAACGAAACCTCGTGCAGCTTGGCGCGTGGCCC 1250638
QY 66 GluAlaSerLysAlaLieserCyLysGlyGlnHisSerIle-----Ser 80
Db 1250637 GAATCCCGCTCTGTGTCAGCGTGCCTGGGAGACCCAGTCGATGCCGCGCATGATGA 1250578
QY 81 TyrThreusSerArg-AenGlnThrValValGlu----- 92
Db 1250577 CAGCATCTGAGCATGCGGCGCATGACCGCGGTGAAACGATCTTGCAGCGCGGCTGC 1250518
QY 93 -----TyrThrisAsp-LysAspThrAspMet-----PheGlnValGlyArgSerThx 109
Db 1250517 CGCGCATCTGCGCCAGACACACAGCGCATTCACACAGATGATGGCGCATCGTCCG 1250458
QY 109 LysErProLleAspPheValValThraPheTrLieserGlySerGlnAsnThraPelu- 128
Db 1250457 CAGGCGCGTCCCATTCACAGACCGGTATGCCAGTGGCGCGGACGATCGAGCGCGT 1250398
QY 129 -----AlaGlnLleThrGlnSerThr- 135
Db 1250397 ATTCGCTCCTCGAATCCTACCTCGCGCGCTTGCCTATTCGACACATCGCGCTCGACGA 1250338
QY 136 -----LieserArgPheAlaCyAspArgLleValCyAspArgAenGluProTyrThra 153
Db 1250337 CCAAGGCGGTACCGCGGTGCGTGCAGGCGATGATTTGGGCAACCTCACCG----- 1250285
QY 153 LaaRgIlePheAlaGlyPheAspSerLysAsnLlePheLeuGly-----GluL 171
Db 1250284 --CGAATCTCG-GGCGCGCGGTGCTGCTTCACCAACCCGATCACCGGACCAACGAC 1250228
QY 171 YsaAlaAlaLysTrpLysAsnProAspArgLysMetAspGlyLeuThrThraPheValL 191
Db 1250227 CGTTCCGCGAATGGCTA-----CCGCGCGCGCAAGCGACAGCTG-----GTGTCCGATTTGG 1250174
QY 191 euValMetHisProArgGlyGlyPheThrgLugIusErGlnProGlyValTrpArgLui 211
Db 1250173 TCGCGGTGCGACGATGACCGCGCTGGCGACAG-----CAACCGCGCTGTCAACGACAC 1250117
QY 211 LeseRValCyGlyAspValLysTrpLysAsnArgGluThraPheSerAlaGlnGlnArgLyl 231
Db 1250116 TG-----ATGCGCGCGCGCGCGCGGATGAGCGCATGCGAG 1250081
QY 231 YsaLeuValGluSerGluThraAsnValLeuGlnAaRgIysEr-----LenuLAspLeuc 249
Db 1250080 GCGACGTTCAGAGATGCAATTCACCGTGGAGAGACGACCACTGTGTGTGACGACCC 1250021
QY 249 YsaGlyAlaThreLeuLeuTrpArgThraAlaAspGlyLeuPheHisThrProThrgLinyuH 269
Db 1250020 GCGGGGCGGAA-----CGCTCGGCGGACGCGCGCGCTGCGCATCTGC 1249970
QY 269 IsrLleGlnAlaLeuArgGlnLuiLAsnAlaAlaArgProGlnCyAspProValGlyLeuA 289
Db 1249969 ATCAGAGGAGACTATGACGACACCGGACACTGCGCGGTGACTCCGACGACATCG 1249910

QY 289 snThreuaAlaPheProSerIle-AaenArgLysGluValValGluGluLysGlnProTyr 308
Db 1249909 AGACTTTGGCTACGCGCGCTGCGTGCAGACCGAAA----- 1249877
QY 309 AlaTyrLeusErCyGlyHisValHisGlyTyrHisAsnTrpGlyHisArgSerAspThr 328
Db 1249876 -----CAGGTTGGCTGCACCGC----- 1249859
QY 329 GluAlaAsnGluArgGluCysPProMetCyAspThrValGlyProTyrValProLeuTyr 348
Db 1249858 ----TTTTGGCCAAAGGCGCTGCGCTTGGCCCGGCGCGGTATTCG----- 1249817
QY 349 LeuGlyCyseGluAlaGlyPheTyrValAspAlaGlyProProThrHisAlaPheThPro 368
Db 1249816 -----GGAACCGCTTACCGGAGTGCAGAG 1249790
QY 369 CyseGlyHisValCyseRgluLysSerAlaLysTyrTrpSerGlnLle 384
Db 1249789 CGCTGACGCTGCGGACCGGCGGACCGGTCAATCTTGTGCGAGATC 1249742
RESULT 11
US-09-103-840A-1/C
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
; US-09-103-840A-1
Alignment Scores:
Pred. No.: 1,7e+04 Length: 4411529
Score: 113.50 Matches: 98
Percent Similarity: 34.0% Conservative: 51
Best Local Similarity: 22.4% Mismatches: 161
Query Match: 5.0% Indels: 129
DB: 3 Gaps: 21
US-10-041-030-4 (1-420) x US-09-103-840A-1 (1-4411529)
QY 4 ProGlyGlnGlu-----GluHieCyAlaProAnlybgluProValLysTyr 19
Db 1251307 CCTGGGCAACAAAGGCGCAGCATGACATGATCGCGGCGACCATCTGCC----- 1251257
QY 20 GlyGluLeuValValLeuGlyTyrAenGlyAlaLeuProAnGlyAaPaRgGlyAaRg 39
Db 1251256 GGTGGCGCGCTGCAATTCGATTACAC-----CGAGGTGGCGGT 1251218
QY 40 LysSerArgPheAlaLeuTyrLysArgProLysAlaAenGlyValLysPro----- 56
Db 1251217 -----GGCTATCTCGCGCGCGCGGATCCAGATCCCGCGCATCGGATGA 1251170
QY 57 -----SerThyValHisValLieserThPro 65
Db 1251169 CGTGTGATCGATGATGATGATGATAAGAAACGAAACCTCGTGCAGCTTGGCGCGTGGCCC 1251110
QY 66 GluAlaSerLysAlaLieserCyLysGlyGlnHisSerIle-----Ser 80
Db 1251109 GAATCCCGCTCTGTGTCAGCGTGCCTGGGAGACCCAGTCGATGCCGCGCATGATGA 1251050


```

Qy      81 TyrThrLeuSerArg-AsnGlnThrValValGlu----- 92
Db      1251049 CAGGATTCGAGATGGGCGCATGACCGAGCGGCTGACGAGTACTTCCAGCGCGGCTGC 1250990
Qy      93 -----TyThrHisAsp-LysAspThrAspMet---PheGlnValGlyArgSerThrG 109
Db      1250989 CGCGGACTTGGCCACGACGACGACGCGGATTCACGACGATGATGCGCGCATCGCGG 1250930
Qy      109 LuserProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu- 128
Db      1250929 CAGCGCGGATCCCATTAACGACGACCGGATGCGGCGGCGGCGGCGGCGGCGGCGG 1250870
Qy      129 -----AlaGlnIleThrGlnSerThr- 135
Db      1250869 ATTGCGCTCTGGAACTCAGCTGCGCGGCTTGCCTATGCGGACCATGACGCGCTTCGACGA 1250810
Qy      136 -----IleSerArgPheAlaCysArgIleValCysAspArgAsnGluProThrTha 153
Db      1250809 CCAAGCGGATACGCGGCGGATGCGAGCGGATGATGCGGAACTCAGCG----- 1250757
Qy      153 LaaArgIlePheAlaAlaGlyPheAspSerSerIlyAsnIlePheLeuGly-----GluL 171
Db      1250756 --CGAACTCG--GGCGCGCGGCGGCTGCTCAGCGCAACCGGATCAGCGGCGGCGGCGG 1250700
Qy      171 yAlaIaAlaLysTrpLysAsnProAspArgIlyIleMetAspArgIleuThrThrAsnGlyValL 191
Db      1250699 CGTTCCGCGGATGCGTA---CCGCGCGCGCAAGCGGCGAGCGAGTGC---GTGTCGCGGATGG 1250646
Qy      191 euValMetHisProArgGlyGlyPheThrGlnGlnSerGlnProGlyValIleThrArgGlu 211
Db      1250645 TCCGCGCGGACCGATACCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1250589
Qy      211 leSerValCysGlyAspValIleThrLeuArgGlnThrArgSerAlaGlnIleArgGlyL 231
Db      1250588 TG-----ATGGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1250553
Qy      231 yAlaValGlnSerGlnThrAsnValLeuGlnAspGlySer-----LeuIleAspLeuc 249
Db      1250552 GCGAGCGTCCAGAGATGGAATTCACCGGATGAGAGACGACGACGCTGCGTGCATGACGACCC 1250493
Qy      249 yAlaGlnIleThrLeuThrArgThrAlaAspGlyLeuPheIleThrProThrGlnLysH 269
Db      1250492 GCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1250442
Qy      269 IsIleGlnAlaLeuArgGlnIleIleAsnAlaAlaArgProGlnCysProValGlyLeuA 289
Db      1250441 ATCAAGAGGAGCTCATGACGACGACGACGACGACGACGCGCGGCGGCGGCGGCGGCGGCGG 1250382
Qy      289 enThrIleuAlaPheProSerIle-AsnArgIleGluValValGluGlyValGlnProTrp 308
Db      1250381 AGACTTTGCTACGCGCGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1250349
Qy      309 AlaTyTrLeuSerCysGlyHisValHisGlyTyThrHisAsnTrpGlnHisArgSerAspThr 328
Db      1250348 -----CAGCGTGGCTGCACCGG----- 1250331
Qy      329 GlnIlaAsnGlnArgGluCysProMetCysArgThrValGlyProGlyValProLeuTrp 348
Db      1250330 ---TTTGGCGCAAGGCGCTGCGCGCTTGC-----CGGCGCGGCGGCGGCGGCGGCGGCGG 1250289
Qy      349 LeuGlyCysGluAlaGlyPheTyValAspAlaGlyProProThrHisAlaPheThrPro 368
Db      1250288 -----GAGCGCGCTACACCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1250262
Qy      369 CysGlyHisValCysSerGluCysSerAlaLysIleTyTrpSerGlnIle 384
Db      1250261 CGTGCAGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1250214

```

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14397
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-14397

Alignment Scores:
Pred. No.: 0.137 Length: 1590
Score: 109.00 Matches: 107
Percent Similarity: 30.1% Conservative: 47
Best Local Similarity: 20.9% Mismatches: 177
Query Match: 4.8% Indels: 182
DB: Gaps: 25

US-10-041-030-4 (1-420) x US-09-252-991A-14397 (1-1590)
Qy      24 ValLeuGlyTyArgGlnAlaLeuProAsnGlyAspArg-----GlyArgArgLys 40
Db      1477 CTGCTGGCGCGGATGGCGGCTGCTCA---GGCGAAAGCGGAGCTGCGGCGGCGGCGGCGG 1421
Qy      41 SerArgPheAlaLeuTyTrpArg-----ProLysAlaAsnGlyValLysProSer 57
Db      1420 CTTCGTCGCGCGCTGCATATCGGTTGCGGAATTCACCTGCCATGATGAGTGGCTGACCGGA 1361
Qy      58 ThrValHisValIleSerThrProGlnAlaSerIlyAlaIleSerCysGlyGlnHis 77
Db      1360 ACTCGCGCGCGCTTCCGCGCATCTCCCGTCCCTGTCGCGGCGGCGGCGGCGGCGGCGGCGG 1301
Qy      78 -----SerIleSerTyThrLeuSerArg----- 85
Db      1300 CCAAGTACCTGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1241
Qy      86 -----Asn 86
Db      1240 CTGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1181
Qy      87 GlnThrValValValGluTyThrHisAspLysAspThrAspMetPheGlnValGlyArg 106
Db      1180 CGATTCGCGCGGCGGATCGGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1121
Qy      107 SerThrGlnSerProIleAspPheValValThrAspThrIleSerGlySerGln-AsnTh 126
Db      1120 ACCGC-AGCGGCTCGATGATTCGTTAC-----GAATACGAAAGGAAAGGAAAGGAAAGGAA 1068
Qy      126 AspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAs 146
Db      1067 TAGCAAGACGCTCACTGAGGCGGCGGATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1031
Qy      146 ArgAsnGlnProTyThrThrAlaArgIlePheAlaAlaGlyPheAspSerSerIlyAsnI 166
Db      1030 -----GGTCAACCGCAATTCACGCGG----- 1008
Qy      166 ePheLeuGlyGlyValAlaLysTrpLysAsnProAspArgIlyIleMetAspGlyLeuTh 186
Db      1007 ATTTCGCGGAACTCGAGGA-----AACCGACGCGTCACTC----- 968
Qy      186 TrpAsnGlyValLeuValMetHisProArgGlyGlyPheThrGlnGlnSerGlnProG 206
Db      967 -----CAACTCG 960
Qy      206 yValTrpArgGlnIleSerValCysGlyAspValIleThrLeuArgGluThrArgSerAl 226

```

RESULT 12
 US-09-252-991A-14397/c
 ; Sequence 14397, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

```

Db      959 TGGCGACGCGGTCTCCG---GCTGGTGTATGCGGACGCTCGGCGAGAACCTTCGC 903
Qy      226 aGInGInArGgLy-LyLeuValGluSerGluThrAnValLeuGlnAsp-----G 243
      902 TGACCATCGAGAGAAAGATG-----GCCATCAGGAGGTGCCAAGATGCCCTCCGCG 849
Qy      243 lSerLeuIlleApleuCyseGly-----A 251
      848 GCCGCGTGGCGGTGATCTGCGGGATCGCCAGATTCAACAGCGCGCGCGAGAGGTGG 789
Db      251 lathrLeuLeuThrArgThr---AlaArgLyLeuPheHleThrPro----- 265
      788 CCAAGCGGTGAGACGGGCGCGCTCGCATCATGTGATGTCGGCGCTGTCTATT 729
Qy      266 -----ThrGlnLysHleGluAlaLeuArgGlnGluLe----- 277
      728 CCTCCAGCGCGACGAGACCGCGGACATTTCCGACGCGTGGCCAGCGGACCGACTGC 669
Qy      278 -----AenAlaAlaArgProGlnCysePro 286
      668 CGGTGATGCTACACACACCGCGCATCTACCGGAGACGTCACCCCGGACATCTCG 609
Qy      286 aGlyLeu-----AenThrLeuAlaPhe-----ProSerIlleAsnArgLys 300
      608 TTTCCCTGGCGGACCTCGAGAACATCGTCTGCTTCAAGACAGCTCCGCGACACCGCC 549
Qy      300 lValValGluGluLysGlnPro-TripAlaTyrlleuSerCyseGlyHleValHleGlyTyr 319
      548 GCTTCATCGACGTACGACACAGGTGGCGGACGCTTCACTCTTTCG---CGGAGCTCG 492
Db      320 HleAsnTrpGlnHleArgSer-----AspThrGlnAlaAsnGluArgGluCysePro 336
      491 ACAGCGGTGCTCGAAGAGCGTGGCGGTGGGTGCCAGGGCTGATCTCGGAAATGTCCA 432
Qy      337 MetCyse----- 338
      431 ACGTGTTCCCGAAGAGGAGGAGACATCTTCGCTGGCCAGCGCGGCGCTTCGCG 372
Db      339 -----ArgThrValGlyProTyrValProLeu 347
      371 AGCGGATCGCGATTCAGAGTGGTGAATGCCGATCTTCATCTGACGCGCGCTCCGACC 312
Qy      348 TrpLeuGlyCyseGluValGlyPheTyrValAspAlaGlyPro----- 361
      311 TGG-----TGACGTGATCATAGCTCGCGAGAACCTGCCGCGCGGCGGACGCGCTCA 258
Db      362 ProThrHleAlaPheThrProCyseGlyHleValCyseSerGluLysSerAlaLysTyrTrp 381
      257 CCGGTCCACCGCGCTGGCTCGCGCGCGCGCGCGATCGCGACAGCATCGAGCGGATCATGG 198
Qy      382 SerGlnLeuProLeuProHleGlyThrHleAlaPheHleAlaLys-----Pro 398
      197 CCAAGG---CCGTGGCGAGCGCGCGCGCGCTGCCGAGCATCGGTGTAGCGGACGCGCC 141
Db      399 PheCyseAlaThrGlnLeuValGlyGln 408
      140 GCCCTGCGCGCTCGCTCCCGCGGAGCGG 111

```

```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14280
; LENGTH: 3801
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14280

Alignment Scores:
Score: 0.562 Length: 3801
Percent Similarity: 30.1% Matches: 107
Best Local Similarity: 20.9% Mismatches: 47
Query Match: 4.8% Indels: 177
DB: 3 Gaps: 182
US-10-041-030-4 (1-420) x US-09-252-991A-14280 (1-3801)

Qy      24 ValLeuGlyTyrAnGlyAlaLeuProAnGlyAspArg-----GlyArgArgLys 40
      802 CTGCTCGCGCGGATGCGGCTGCTCA---GGCGAAGGCCAGCTCGGGTCCGGAACG 858
Db      41 SerArgPheAlaLeuTyrLysArg-----ProValAlaAnGlyValLysProSer 57
      859 CTTCGTGGCGGCTCCATCATCTGTTGGAATTCACCTGCCATGAGATGCTGACCGA 918
Qy      58 ThrValHleValIlleSerThrProGlnAlaSerLysAlaIlleSerCyseGlyGlnHle 77
      919 ACTCGCGCGCGCTTCGGCATCTCCCGTCCCTGTCGGCGGCGCGCTGATCACCGGANT 978
Db      78 -----SerIlleSerTyrThrLeuSerArg----- 85
      979 CCACAGTACTGCTCGCGCGCGGACGATCTCGCGCGAAGCTATCGCTGCCGAC 1038
Qy      86 -----Asn 86
      1039 CTGCGCGCGGCGCTGCTGACCAACCGCGGACGAGCGCGCGCGGAGAAC 1098
Db      87 GlnThrValValValGlyTyrThrHleAspLysAspThrAspMetPheGlnValGlyArg 106
      1099 CGATCCGCGCGCGATCGGATCGCCACGACGAGGAGAACGCTCGCGCTGTGACTCGG 1158
Qy      107 SerThrLysSerProIlleAspPheValValThrAspThrIlleSerGlySerGln-AsnTh 126
      1159 ACCGC-AGCGGTCCGATGTAATTCGTATAC-----GAATATGAAAGGAGAACCA 1211
Db      126 rAspGluAlaGlnIlleThrGlnSerThrIlleSerArgPheAlaCyseArgLysVal 146
      1212 TGAAGAACGACGTCACTGAGAGCGGCTATTCGCGC----- 1248
Qy      146 rArgAnGluProTyrThrAlaArgIllePheAlaAlaGlyPheAspSerSerLysAnIl 166
      1249 -----GTTCAACACCAATTCAAGCGG 1271
Db      166 ePheLeuGlyLysLysAlaAlaLysTyrLysAnProAspGlyHleMetAspGlyLeuTh 186
      1272 ATTTCTCGTGAACCTGAGGA-----AACCCACAGGTCATCTC----- 1311
Qy      186 rThrAnGlyValLeuValMetHleProArgGlyGlyPheThrGluGluSerGlnProGln 206
      1312 -----CAACCTGG 1319
Db      206 ValTrpArgGluIlleSerValCyseGlyAspValTyrThrLeuArgGluThrArgSerAl 226
      1320 TGGCGACGCGGTCTCCG---GCTGTGTATTCGCGGACGCTCGCGAGAACACTTCGC 1376
Qy      226 aGInGInArGgLy-LyLeuValGluSerGluThrAnValLeuGlnAsp-----G 243
      1377 TGACCATCGAGAGAAAGATG-----GCCGTACGAGGTTCGCAAGATGCTCCGCGG 1430
Db      243 lSerLeuIlleApleuCyseGly-----A 251
      1431 GCGCGTGGCGGTGATCTGCGGGATCGCGAGTTCAACGCGCGCGCGGAGAGTGG 1490

```

```

Qy 251 lathrleuLeuTprargThr---AlaapglLeuPheHistrPro----- 265
Db 1491 CCAAGCGGTGGAGCGGTGGCGTGCAGCGCATTCATGTGATGCGCGCTGTTATT 1550
Qy 266 -----ThrGlnyHsiIleGluAlaLeuArgGlnGluIle----- 277
Db 1551 CCTCAAGCGCGACAGACCGCGAGCATTTCCGACGCGTGGCGGACGCGACGACTGC 1610
Qy 278 -----AsnAlaAlaArgProGlnCyseProy 286
Db 1611 CGGTGATGTCTACAACAACCGCGCATCTACCGGAAACGACGTACCCCGACATCTGG 1670
Qy 286 alGlyLeu-----AsnThrleuAlaPhe-----ProSerIleAsnArgIys 300
Db 1671 TTTCCTGGCGGACCTGCGAGAACATCTGTCTTCAAGACACTCCGCGGACACCCGCC 1730
Qy 300 luvAlaIgluGluYsglnPro-TPrAlaTyrlLeuSerCyseGlnIvalHIsGlyTr 319
Db 1731 GCTTCATCGACGACGACGACGACGAGTGGCGAGCGCTTCACTCTTCG---CCGCGCTCG 1787
Qy 320 HIsAsnTrpGlnHIsArgSer-----AspThrGluAlaAsnGluArgIuCysePro 336
Db 1788 ACAGACGTGTCTGAAAGCGTGCAGGTGGGTGCCAGCGCTGAGATCTCGGAAATGCCA 1847
Qy 337 MetCys----- 338
Db 1848 ACSTGTTCGCGAAGAAAGCGGACATCTTCGCGCTGGCGGAGCGCGCTTGGCGG 1907
Qy 339 -----ArgThrValGlnProIleu 347
Db 1908 AGCGGATGCCGATCTACGAGTGGGTATGCCGATCTCCATCTCGACGCCCGCTCCGACC 1967
Qy 348 TrpLeuGlnCyseGluAlaGlyPheTyrlValAspAlaGlyPro----- 361
Db 1968 TGG-----TGCAGTGCATCAAGTCTGCAGACACTCGCGGTTCGCGGACGCGCTCA 2021
Qy 362 ProThrHIsAlaPheThrProCyseGlnIvalCyseSerGluYsserAlaIlyTrp 381
Db 2022 CCGGTCCACGCGCGCTGGCGCTCGCGCGCGCGCATCGCGACGACGTGACGAGTATGG 2081
Qy 382 SerGlnIleProLeuProHIsGlyThrHIsAlaPheHIsAlaIlyCys-----Pro 398
Db 2082 CCAAGC---CCTGGCGACCGCGCGCGCGCGCTCGCGACATCGGCTTTCAGCCAGCGCCC 2138
Qy 399 PheCysAlaThrGlnLeuValIgluGln 408
Db 2139 GCCCTCGCGCGCTCGCTCCCGCGGAGCGG 2168

RESULT 14
US-09-949-016-12345/c
; Sequence 12345, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12345
; LENGTH: 86414
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (1)...(86414)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12345

Alignment Scores:
Pred. No.: 147
Score: 107.00
Percent Similarity: 33.3%
Best Local Similarity: 21.9%
Query Match: 4.7%
DB: 3
Gaps: 26

US-10-041-030-4 (1-420) x US-09-949-016-12345 (1-86414)

Qy 4 ProGlnGlnGluHIsCyseAlaProAsnIyGluProValIyTrGlyLeuVal 23
Db 33554 CCGGCGACAGCGCGACCGCGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 33516
Qy 24 ValLeuGlyTyrlAsnGlyAlaLeuProAsnIyAspArgGlyIyArgIySerArgPhe 43
Db 33515 -----CAAGAGGCTCAGCTCCCTAAGC----- 33492
Qy 44 AlaLeuTyrlAspArgPro---LysAlaAsnGlyValIyProSerThrValHIsValIle 62
Db 33491 AGCAGCTACCCCTCGCGCACAGAAAGTACAGGACATTTCCAGACCAACCTAATGAA 33432
Qy 63 SerThrProGlnAla-SerLysAlaIleSerCyseGlnIvalHIsSerIleSerTyrlH 82
Db 33431 GTTGTTTTAAAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 33372
Qy 82 rIeuserArgAsnGlnThrValValIyGluTyrlHIsAspIys----- 97
Db 33371 TCTCTTTCACACACACACACACACACACACACACACACACACACACACACACAC 33312
Qy 98 -----AspThrAspMetPheGln-- 103
Db 33311 AGTATGATCTGACCCACTTATGCTGGTGAAGCTGGAGAGGAGTACTATTCAGGC 33252
Qy 104 -----ValGlyArgSerThrGlnIleAspPheValIva 116
Db 33251 CTGATCATCAGCGCACAGGCGGATGGGTGGAGGAGCGGATGGGTGATACCTTC 33192
Qy 116 ThrAspThrIleSerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrI 136
Db 33191 GCTGCTCCCTCGAGTGCCTCGGAAAGCAGAGTGCCTCTCTGATGAAAGCCACT 33132
Qy 136 eSerArgPheAlaCyseArgIleValCyseAspArgAsnIyProTyrlH-AlaArgIleP 156
Db 33131 CAGAACATGCTCTGCGCGCTC-----TGCATTCCTCCCTGCGCGACTCT 33087
Qy 156 heAlaAlaGlyPheAspSerSerLysAsnIlePheLeuGlyGluValAlaIyIys---- 174
Db 33086 GCGGCGGTGGGAGACATCAGACGCGCAG-----CTGGAGCCCAATACACCGAGACG 33033
Qy 175 -----TrpLysAsnProAspGlyHIsM 182
Db 33032 CACCCACTGCGCTTCAGGACGCGCGCAACGAGATTTGTTGGAAACACCCAGCGGAC 32973
Qy 182 eAspGlyLeuThrThrAsnGlyValIleuValMechIAspPro-----ArgGlyG 198
Db 32972 TCAGTGGGCTC---GTTGGGAAATCCAAATCAGTGGCGCTCTGCAAGAGCAAGAA 32916
Qy 198 lYpHeThrGlnGluSer-----GlnProGlyValIyTrp 208
Db 32915 GG-----GATGCACGCTTCCCTCGCGAGATCCAGAGCCCTGGCGAGGTGGCT 32862
Qy 209 -----ArgGluIleSerY 213
Db 32861 GCCAGGTCCCTCTCACTGTGATGCCATCTTGAATGATAGACGACGACTCAGACA 32802
Qy 213 alCyseGlyAspValTyrlThrLeu-----ArgGluThr-----ArgS 225
Db 32801 TCAGCGGTCTCTCTTCAAGTTTGAAGAGTGTGTGTCTGCAGGAGACAGAGATCAT 32742

```

QY 225 ealaelnglnlArGlyLys-----LeuValGluSerGluThrAsnVal----- 239
DB 32741 CTCTCTTCGAGAGGGGAGAGAGCTGTATGTCGGCAGACACCCCTCCAACTCTCTCCAG 32682
QY 240 -----LeuGlnAerGlySerLeu-----IleAspLeuCyGlyValAthrLeuLeuTrpA 256
DB 32681 GCACAGACTACACACAGAGGCTCCCTGCCCCGCACTTGAGCTGGTGCCACAGGCTG----- 32626
QY 256 rghrhlAerGlyLeuPheHieThrProthrglnlyvniIleGluAlaLeuArg----- 274
DB 32625 -----GTCATCTCAAGCCAGAAATGAGAAACACAGAAAGCTCA 32586
QY 275 -----GlnGluIleAsnAlaAthrProGlnIleCyPProValGlyLeuAthr 290
DB 32585 AGCCGAGGCCCGACCCACCAACAGCAGCGCGCGGGAGACCCGACACCCCGCATTCAC 32526
QY 291 -----LeuAlaPheProSerIleAsnArgLyGluValGluGluLeuGlnP 307
DB 32525 CCAAGCTGGCAGACAGCCCGACAGCTGACAGAGAGAGAGAGCTTCAGAGAGGAGAG 32466
QY 307 rofrrAlaTrpLeuSerCyGlyvniValHieGlyTrpHieAthrGlyHieArgSerA 327
DB 32465 AGCCCGAGCTCTGAGCATGTGGGGCCCTGGGGCCCTGAGGCTGGGGT----- 32416
QY 327 eprhrGluAlaAerGlyAerGlyCyPProMetCyAthrGlnValGlyProTrpValAthr 347
DB 32415 -----GAGCCCCAGATGTCAGACCCAGAGGCTTCAGAGAGAGAGAGAGAG 32380
QY 347 euTrp-----LeuGlyCyGlyValAerGlyPheTrpValAthrAerGlyP 361
DB 32379 --TGGGGTCTGGGGTCCAGAGCTTGGAGCTGTAATAGGGGCTC---TAGGAGCCCATTC 32325
QY 361 roProthrglnAlaPheThrProCyGlyvniValCySerGlyLysSerAlaLys 379
DB 32324 CAGCTCAGACAGCCAGAAACATGGGCTCAGCTGGAGAGCCAGAGAGAGAGAG 32269

RESULT 15
US-09-949-016-15758/c
Sequence 15758, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTRI, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15758
LENGTH: 86414
TYPE: DNA
ORGANISM: Human
FEATURES:
NAME/KEY: misc_feature
LOCATION: (1)_(86414)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15758

Alignment Scores:
Pred. No.: 147
Score: 107.00
Percent Similarity: 33.3%
Best Local Similarity: 21.9%
Query Match: 4.7%
DB: 3
Length: 86414
Matches: 105
Conservative: 55
Mismatch: 166
Indels: 154
Gaps: 26

US-10-041-030-4 (1-420) x US-09-949-016-15758 (1-86414)
QY 4 ProGlyGlnGluGlnIleCyValAerGlnProVallyGlyGlyGlyLeuVal 23
DB 33554 CCGGCGAGAGAGCCAGCTGTGGCCCTCCAGCCCT----- 33516
QY 24 ValLeuGlyTrpAerGlyAlaLeuProAerGlyAerArgGlyAerArgLysAerArgPhe 43
DB 33515 -----CAAGAGGCTCAGCTCCCTTAAGC----- 33492
QY 44 AlaLeuTrpValAerPro-----IleValAerGlyValAerProSerThrValHieValIle 62
DB 33491 AGCAGCTACCCCTGGCCAGAGAAATGACAGAGGATTTCCAAAGACCAAGTAAGAA 33432
QY 63 SerThrProGlnAla-SerValAlaIleSerCyValGlyGlnHieSerIleSerTrpH 82
DB 33431 GTTGTTTTAAAGGTGGCCCAATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 33372
QY 82 rLeuSerArgAerGlnIleThrValValAlaGluTrpHieAerLys----- 97
DB 33371 TCTCTCTTCAC 33312
QY 98 -----AspThrAerPheGln----- 103
DB 33311 AGTATGACTGTGACCCCATTAAGTGTGGGTAAAGTGGAGAGGAGTATTCAGAGC 33252
QY 104 -----ValGlyAerSerThrGlySerProIleAerPheValA 116
DB 33251 CTGATGATCAGCAGCAGAGGAGCATGGGTGGGTAGAGAGAGAGAGAGAGAGAGAGAGAG 33192
QY 116 lThrAerThrIleSerGlySerGlnAerThrAerGlyAlaGlnIleThrGlnSerThrI 136
DB 33191 GCTGGCTCCCTGAGAGTGGTCCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 33132
QY 136 eSerArgPheAlaCyAerGlyLeuValCyAerArgAerGlnProTrpTrpAlaArgIle 156
DB 33131 CAGAACATGCTCTGGCGGCTC-----TGGCATCTCTCTCTCTCTCTCTCTCTCTCTCT 33087
QY 156 heAlaAlaGlyPheAerSerSerLysAerIlePheLeuGlyGlyValAlaLys----- 174
DB 33086 GCGGGGTGGGGGAGCAGCATGACAGCGGAG-----CTGGAGCCCAATACACACAGAGAG 33033
QY 175 -----TrpLysAerProAerGlyHie 182
DB 33032 CAGCCAGCTGCTCCAGAGAGCCCAACAGAGATTTGGGAAACACAGAGAGAGAGAGAG 32973
QY 182 eAerGlyLeuThrAerGlyValLeuValAerMetAerPro-----ArgGly 198
DB 32972 TCAGTGGGGTTC---GCTGGGAGAGTCAAACTGAGTGGCCCTTCAGAGAGAGAGAGAG 32916
QY 198 lPheThrGlyGluSer-----GlnProGlyValTrp----- 208
DB 32915 GG-----GAGTCAAGCT 32862
QY 209 -----ArgGluIleSerV 213
DB 32861 GCGAGTCCCTCTCACTGATGATCCATCTTGAATGATGAGAGAGAGAGAGAGAGAGAGAG 32802
QY 213 aLcyGlyAerValTrpThrLeu-----ArgGluThr-----ArgS 225
DB 32801 TCAGCGGGTCTCTCTCAAGTTGACAGAGGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAG 32742
QY 225 ealaelnglnlArGlyLys-----LeuValGluSerGluThrAsnVal----- 239
DB 32741 CTCTCTTCGAGAGGGGAGAGAGCTGTATGTCGGCAGACACCCCTCCAACTCTCTCCAG 32682

QY 240 -----LeuGlnAerGlySerLeu-----IleAspLeuCyGlyValAthrLeuLeuTrpA 256
DB 32681 GCACAGACTACACACAGAGGCTCCCTGCCCCGCACTTGAGCTGGTGCCACAGGCTG----- 32626
QY 256 rghrhlAerGlyLeuPheHieThrProthrglnlyvniIleGluAlaLeuArg----- 274

this page blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: March 2, 2006, 05:14:56 ; Search time 5360 Seconds
(without alignments)
3666.152 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MRSPOGHEHCAPIKPEVKYK.....ATQVGEQNCIKLIFQSPID 420

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/abs/ABSWEB.spool/US10041030/runat_01032006_134406_21863/app.query.fasta_1
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MIMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs808
-USER=US10041030 @CGM.1.1 5315 @runat_01032006_134406_21863 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *
10: gb_est10: *
11: gb_est11: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2151.5	94.0	3387	4 AK030564	Mus muscu
2	2147	93.8	1186	10 AY409116	AY409116 Homo sapi
3	2052.5	89.6	1183	10 AY409118	AY409118 Mus muscu
4	1926	84.1	3530	4 AK045515	AK045515 Mus muscu
5	1912	83.5	2584	4 AK045673	AK045673 Mus muscu
6	1681	73.4	1186	10 AY409117	AY409117 Pan trogl
7	1632	71.3	1879	4 AK044418	AK044418 Mus muscu

8	1626.5	71.0	1242	10 AY408790	AY408790 Pan trogl
9	1622	70.8	2572	4 CR606152	CR606152 full-length
10	1621.5	70.8	1242	10 AY408789	AY408789 Homo sapi
11	1620	70.7	2710	4 HSM805479	AL834395 Homo sapi
12	1611.5	70.4	1242	10 AY408791	AY408791 Mus muscu
13	1563	68.3	2474	4 AK033815	AK033815 Mus muscu
14	1519	66.3	849	5 BX388547	BX388547 BX388547
15	1495	65.3	855	8 CX975821	CX975821 JGI CAAP9
16	1473	64.3	3531	4 AK029586	AK029586 Mus muscu
17	1430.5	62.5	957	7 CN329888	CN329888 AGENCOURT
18	1391	60.7	869	5 BX688136	BX688136 BX688136
19	1363.5	59.5	856	6 CF241089	CF241089 AGENCOURT
20	1362	59.5	1117	7 CK027873	CK027873 AGENCOURT
21	1352	59.0	829	8 CX314078	CX314078 JGI XZT96
22	1344.5	58.7	878	8 CX941650	CX941650 JGI CAAP7
23	1334.5	58.3	1076	3 BM800644	BM800644 AGENCOURT
24	1314	57.4	809	5 BX700555	BX700555 BX700555
25	1306.5	57.1	764	3 BP682660	BP682660 BP682660
26	1294	56.5	701	7 CK637810	CK637810 UI-M-HO0
27	1261	55.1	768	7 CO738704	CO738704 SLIB04C21
28	1259.5	55.0	726	3 BP701801	BP701801 BP701801
29	1258	54.9	784	5 BX749714	BX749714 BX749714
30	1237	54.0	1229	4 BC026695	BC026695 Mus muscu
31	1207	52.7	741	8 CX744549	CX744549 JGI ANBT2
32	1203.5	52.6	752	3 BP709958	BP709958 BP709958
33	1192.5	52.1	1188	8 DN729302	DN729302 CMB46-E04
34	1189	51.9	747	5 BX774759	BX774759 BX774759
35	1187.5	51.9	773	8 CX439604	CX439604 JGI XZG77
36	1187	51.8	738	8 CX941649	CX941649 JGI CAAP7
37	1187	51.7	752	8 CX443252	CX443252 JGI XZG89
38	1185	51.7	1006	3 BO049269	BO049269 AGENCOURT
39	1183	51.7	816	7 CR975985	CR975985 CR975985
40	1169	51.0	732	8 CX436008	CX436008 JGI XZG58
41	1161	50.7	766	6 CF539872	CF539872 UI-M-BX0
42	1159	50.6	889	9 BX327874	BX327874 BX327874
43	1155.5	50.5	832	8 DN521048	DN521048 1263866 M
44	1154.5	50.4	795	8 CK638809	CK638809 UI-M-HO0
45	1148	50.1	763	5 BM959532	BM959532 BM959532

ALIGNMENTS

RESULT 1	AK030564	3387 bp	mRNA	linear	HTC 03-Apr-2004
LOCUS	AK030564				
DEFINITION	Mus musculus adult male pituitary gland cDNA, RIKEN full-length enriched library, clone:530432g13 product:pellino 2, full insert sequence.				
ACCESSION	AK030564	GI:26326560			
VERSION	AK030564.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
1	Carninci, P. and Hayashizaki, Y.				
2	High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
10349636					
2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komu, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
3	Genome Res. 10 (10), 1617-1630 (2000)				
11042159					
3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komu, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,				

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 3387) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komori, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

1. 3387

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:5330432G13"
/db_xref="taxon:10090"
/clone="5330432G13"
/sex="male"
/tissue_type="pituitary gland"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
336..1595
/note="unnamed protein product; peliino 2 (MGD|MG1:1891445, GB|NM_033602, evidence: BLASTN, 99%, match=864)
putative"
/codon_start=1
/protein_id="BAC27024.1"
/db_xref="GI:26326561"
/translation="MSPSGKPEPTPKKPKVKKRELIVYGYNGLPVGDRGKSRPA LYKRTASGVKPESTIHNVSTPQASKAISGSHSISTTSRSQTVVETTHDQDTMF QVNRSTESPLIDFVDTVSGQNDAGITOSTISIPKFCRLVCRNREYTKRIAPAGD SSQKIFLGEKRAKMDKPDGMDGLTNGVLVMMHPOGFTESQPGVREISVCGDVT LRETRAQGRGLVESRTNVLQDGLDLCATLMTADGLFPAPTOKHIALROBI MAAPQCPVGLNTLAPPSINRKEVVEKQWAVYLSGVHGVSHWGRSDTEANERBC

ORIGIN
Alignment Scores:
Pred. No.: 1,92e-205 Length: 3387
Score: 2151.50 Matches: 395
Percent Similarity: 96.7% Conservative: 11
Best Local Similarity: 94.0% Mismatches: 13
Query Match: 94.0% Indels: 1
DB: 4 Gaps: 1

US-10-041-030-4 (1-420) x AK030564 (1-3387)

Qy 1 MetPheSerProGlyInGluGluHisCyAlaProAsnLysGluProValLysTyrGly 20
Db ATGTTTTCCTCCCGGGCCANNAAGAACCCANCAACCCCAAGGAGCAGGAAATCCGG 395

Qy 21 GluLeuValValLeuGlyTyrAenglyAlaLeuProAsnGlyAspArgGlyArgTyrGly 40
Db GAGCTGTGGTCTCTGGGTACATGCTGTCTTACCTAATGTCACAGGGGAGGAGGAA 455

Qy 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db AGCAGATTGGCCCTCTAATAAGCGACCTACGCGGTGTGTCAACCCAGCAATCCAC 515

Qy 61 ValLysSerThrProGlnLysSerLysAlaLysSerCyAlaLysGlyInHisSerLys 80
Db ATGCTTCCACACACAGGCGGTCCAGGCTCATGCTCCAGAGACATCCAGCATATCG 575

Qy 81 TyrThrLysSerArgAsnGlnThrValValValGluTyrThrHisAspLysAspThrAsp 100
Db TACAGTTGTCACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 635

Qy 101 MetPheGlnValGlyArgSerThrGluSerProLysAspPheValValThrAspThrLe 120
Db ATGTTTCAGGTGGGAGGTGTCACAGAAAGCCCATGATGCTGTGTGTACAGACAGGTT 695

Qy 121 SerGlySerGlnAsnThrAspGluAlaGlnLysThrGlnSerThrLysSerArgPheAla 140
Db TCCGGCGGTGACAGCAAGT--GCCAGATACACAGACCCATCTCTAGTTGCA 752

Qy 141 CyArgIleValCyAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db TGCAGATCTGTGTGTACAGAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 812

Qy 161 AspSerSerLysAsnLysPheLeuGlyGluLysAlaAlaLysTyrLysAsnProAspGly 180
Db GATTTCTCCAAATATATCTTCTTGTGAGAGAAAGCAGCAAAATGGAACCACTGATGGA 872

Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db CACATGATGATGATCTCTACCAACAGGTGTCGTGTATGACCCGGAAGGAGCTTCACC 932

Qy 201 GluGlySerGlnProGlyValTPAArgGluLysSerValCyArgIleAspValTyrThrLeu 220
Db GAGGAATCCAGCCTGAGATCTGAGAGAGATCTCTGTGTGGAGATGATACACTTG 992

Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
Db CGAGACACACAGTCGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1052

Qy 241 GlnAspGlySerLeuLysPheLeuCyArgIleAlaThrLeuLeuThrPheThrHisAspGly 260
Db CAAGACGGCTCCCTCATGACCTGTGTGGGACCACTCTCTGAGAAACCGAGATGGC 1112

Qy 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnLysIleAsnAlaAla 280
Db CTTTTCACGCTCTTACTCAGAGCAGCATGAGACCTCCGAGAGAGATCAATGACGCC 1172

Qy 281 ArgProGlnCyAspProValGlyLeuAsnThrLeuAlaPheProSerLysAsnArgLysGlu 300
Db CGACCCAGTGGCCGAGGGGCTTAAACCTGTGCTTCCCGAGATCAACCGAGAGAA 1232

QY 301 ValValGluGluLysGlnProTTPAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
 Db 1233 GTGGTGGAAGAAGACGACCCCTGGGCACTGAGCTGGCGGCAATGCGAGGCTACAC 1292
 QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
 Db 1293 AGCTGGGGCCATGCGACCGACGAGAGCCAGAGAGGAGGTGTCCCATGTGCGAGACT 1352
 QY 341 ValGlyProTyrValProLeuTTPLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
 Db 1353 GTGGGCCCCACGCTCTCTGTGGCTGGCTGTGAGGACAGATTATATGTCATGGCGGA 1412
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
 Db 1413 CCCCACACTCAGCGTTTCACCCCTGGCGGACGCTGTTCAGAGAGATGTCGCAAGTAC 1472
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
 Db 1473 TGGTTCGAGATCCCATGCGCCCGACGAGACGACGCGTTTCACGCCGCTGTCTCTGC 1532
 QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
 Db 1533 GCCACGCGAGCTGGTGTGTAACAGAACTGCATCAATGATTTTCCAGAGTCGAGTGAC 1592
 RESULT 2
 AY409116 1186 bp DNA linear GSS 16-DEC-2003
 LOCUS Homo sapiens PBL12 gene, VIRtual TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY409116 GI:39765084
 VERSION AY409116.1
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.
 REFERENCE 1 (bases 1 to 1186)
 Clarke, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
 Todd, M.A., Tanenbaum, D.M., Civejlo, D.R., Lu, F., Murphy, B.,
 Portiera, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J.,
 Adams, M.D. and Cargill, M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1186)
 Clarke, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
 Todd, M.A., Tanenbaum, D.M., Civejlo, D.R., Lu, F., Murphy, B.,
 Portiera, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J.,
 Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1. 1186
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>1186
 /gene="PBL12"
 /locus_tag="HCM3450"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,1e-205 Length: 1186
 Score: 2147.00 Matches: 394
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 93.8% Indels: 0

DB: 10 Gaps: 0
 US-10-041-030-4 (1-420) x AY409116 (1-1186)
 QY 27 TYZAsnGlyAlaLeuProAsnGlyAspArgGlyArgGlyLysSerArgPheAlaLeuTyr 46
 Db 2 TACAAATGGCTTATCCCAATGAGATAGAGGACGAGAGAAAGATGATTCCTAC 61
 QY 47 LysArgProLysAlaAsnGlyValLysProSerThrValHisValIleSerThrProGln 66
 Db 62 AAGCGGCCAAGGCAATGATGTTCNAACCCAGACCGCTCATGTGATTCACGCCCCAG 121
 QY 67 AlaSerValAlaIleSerCysLysGlyGlnHisSerIleSerTyrThrLeuSerArgAsn 86
 Db 122 GCATCCAGGCTATCAGCTGCAAGGTCACACAGATATCTTCACTTTGTCAAGGAAT 181
 QY 87 GlnThrValValAlaGluTyrThrHisAspLysAspThrAspMetPheGlnValArg 106
 Db 182 CAGACTGTGGTGTGAGTACACATGATAGATACGATACGATGTTTCAGGTGGCACA 241
 QY 107 SerThrGluSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThr 126
 Db 242 TCACAGAAAGCCCTATTCATCTTGTGCACAGACATATTCGACAGCCACGACG 301
 QY 127 AspGluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAsp 146
 Db 302 GACGAGGCCAGATCAACAGACACATATCCAGGTGCTGCTGACGATGTGTGAC 361
 QY 147 ArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPheAspSerSerLysAsnIle 166
 Db 362 AGGAATGAACTTACACAGACGAGATATTCGCCGCGGATTGACTTCCAAAACATA 421
 QY 167 PheLeuGlyGluValAlaAlaLysTyrLysAspProAspGlyHisMetAspGlyLeuThr 186
 Db 422 TTCTTGGAGAAAGGACAGCAAGATGGAAGAAACCCCGACGCGCACATGATGGCTCACT 481
 QY 187 ThrAsnGlyValLeuValMetHisProArgGlyGlyPheThrGluGluSerGlnProGly 206
 Db 482 ACTAATGGCTCTGTGTGATGATCCACAGAGGGGCTTCAACGAGAGTCCAGCCGGG 541
 QY 207 ValTrpArgGluIleSerValCysGlyAspValTyrThrLeuArgGluThrArgSerAla 226
 Db 542 GTTCGGCGGAGATCTCTGTCTGTGAGATGTATCACTTGGAGAAACAGTGGCC 601
 QY 227 GlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeuGlnAspGlySerLeuIle 246
 Db 602 CAGCAACGAGAAAGCTGGTGAAGAGAACCAACCTCTGACAGACGCTCCATCAT 661
 QY 247 AspLeuCysGlyAlaThrLeuLeuTTPArgThrAlaAspGlyLeuPheHisThrProThr 266
 Db 662 GACCTGTGGGGCCACTCTCTCTGAGAAACGAGATGGGCTTTTCACTATCCAACT 721
 QY 267 GlnLysHisIleGlnAlaLeuArgGlnGlnIleAsnAlaAlaArgProGlnCysProVal 286
 Db 722 CAGAGACATTAAGAGCTCCCGGACGAGATTAAGCCCGGCTCAGTCTCTGTG 781
 QY 287 GlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValAlaGluGln 306
 Db 782 GGGCTCAACACCTGGGCTTCCCAAGATCAACAGAAAGAGTGTGTGAGAGAAAGCAG 841
 QY 307 ProTTPAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsnTrpGlyHisArgSer 326
 Db 842 CCTGGGCAATATTCAGTGTGTGCGACGTCAGAGGATCAACAATCGGGCCATCGAGT 901
 QY 327 AspThrGluAlaAsnGluArgGluCysProMetCysArgThrValGlyProTyrValPro 346
 Db 902 GACACGAGGCCAACAAGAGAGAGGTGTCCATGTGACAGACTGTGGCCCTTATGTGCT 961
 QY 347 LeuTrpLeuGlyCysGluAlaGlyPheTyrValAspAlaGlyProProThrHisAlaPhe 366
 Db 962 CTCGTGGCTTGTGTGAGGACGATTTTATGACGACGAGACCGCAACTCATGTGCTTTC 1021
 QY 367 ThrProCysGlyHisValCysSerGluLysSerAlaLysTyrTrpSerGlnIleProLeu 386

Db	1022	ACTCCCTGTGGACACAGTGTGCTCGGAGAAAGTCTGCMAAATACTGCTTCACATCCCGTTG	1081
Qy	387	ProHsiGlyThrHisAlaPheHisAlaAlaCysProPheCysAlaThrGlnLeuValGly	406
Db	1082	CCTCATGGACACTGATGCATTTCCAGCGCTCTGGCCCTTTCTGTGCTACACAGCTGTTGGG	1141
Qy	407	GluGlnGlyCysGlyLeuValLeuLeuLeuPheGlnGlnGlyProLeuLeu	420
Db	1142	GAGCAAAACTGCATCAATAATTAATTTCCAAAGTCCCAATTGAC	1183
RESULT 3			
AY409118			
LOCUS			
DEFINITION	AY409118	1183 bp	DNA linear GSS 16-DEC-2003
ACCESSION	Mus musculus PEL12 gene, VIRUAL TRANSCRIPT, partial sequence,		
VERSION	AY409118		
KEYWORDS	AY409118.1 GI:39765086		
SOURCE	GSS:		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 1183)		
TITLE	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 1183)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..1183		
gene	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
	<1..>1183		
	/gene="PEL12"		
	/locus_tag="HMC3450"		
ORIGIN			
Alignment Scores:			
Score No.:	3.84e-196	Length:	1183
Score:	2052.50	Matches:	375
Percent Similarity:	97.7%	Conservative:	10
Best Local Similarity:	95.2%	Mismatch:	8
Query Match:	89.6%	Indels:	1
DB:	10	Gaps:	1
US-10-041-030-4 (1-420) x AY409118 (1-1183)			
Qy	27	TyrAsnGlyAlaLeuProAsnGlyValAspArgGlyArgGlySerArgPheAlaLeuTyr	46
Db	2	TACATAGTGTCTTTCCTATATGCTGACAGGGGACAGAGAAAGCAGATTGGCCCTCAT	61
Qy	47	LysArgProLysAlaAsnGlyValIleProSerThrValHisValIleSerThrProGln	66
Db	62	AAGCGAGCCTTACGCGACGATGTCTCAACCCAGCAACATCCCATGCTTCCACACACAG	121
Qy	67	AlaSerLysAlaIleSerCysValSerGlyGlnHisSerIleSerTyrThrLeuSerArgAsn	86

Db	122	GGCTTCAAAGGCATCAGCTCCAGAGCATCAGAGCATATCTGACGTTTCACGGAGC	181
Qy	87	GIInThValValValGIuTYrThHiAePrlYaePThrAePMePheGIInValGIYArG	106
Db	182	CAGACGGTAGTGGTGGAGTACACACAGATTAAGACACGGCAGATTTCAAGTGGGACAG	241
Qy	107	SeThrGIuSePrcIIeAePheValValThrAePThrIIeSePrlYsePrlInAePThr	126
Db	242	TCACAGAAAGCCCCATTGACTTGTGTGTCTACACACAGGTTTCGGGGCTCAAGAACAA	301
Qy	127	AsPrluAInaGIInIleThrGIInSeThrIIeSePrlYpheaAIIeValCYsAglIeValCYsAeP	146
Db	302	GAT---GCCAGATACACACAGAGCACCATCTCTAGTTCGATTCAGATGCTGTGTGAC	358
Qy	147	ArgAeGIuPrcIYrThrValaYrIIeAePheAIIaGIYpheaPSeSePrlYaePrlIe	166
Db	359	AGGAACAGACCATGTATACGACACGCATATTCGGCGACAGATTGATTTCCAAAATATTC	418
Qy	167	PheIneGIYGIuYsValaAlaYsPrlYsAePrcIYpheaPrlYhIsmcAePrlYleuThr	186
Db	419	TTTTCTGGAGAAAGACGACAAATATGAAAAACCTTATGACACATGAGATGACTACT	478
Qy	187	ThrAeGIYValIeValIeValIeMeuHIePrcIYpheaAIIeGIYpheaThrGIuSePrlY	206
Db	479	ACCAATGCTGTCTAGTATGATGCACCCGACAGAGAGCTTACCCAGAGAAATCCACCTGGA	538
Qy	207	ValTPraPrgIuIIeSePrlCYeGIYAsPrlYrThrIleuAArgIuThrArGserAIIe	226
Db	539	GTCGTGAGAGAGATCTGTCTGTGGGGATGTGTACACTTGCAGAGACACAGGTCTGGCC	598
Qy	227	GIInIInPrgIYrYsIeValGIuSePrlYhIsmcAIIeGIInAePrlYsePrlIe	246
Db	599	CAGACAGAGGAAAGCTGTGTGGAAAGTGAACCAACGCTCTGCAAGACGGCTCTCATTT	658
Qy	247	AsPrluCYsGIYAThThrIleuThrTPraPrgThrValaAsPrlYleuPheIsthPrcIY	266
Db	659	GACTGTGTGGGGCACTCTCTGTGAGAAACGCAATGGCTTTTTCAGCTCTACT	718
Qy	267	GIInYsIIeGIuAIIeAePrgIInGIuIIeAePrlYsGIuAIIaAArgPrcIInCYsPrcI	286
Db	719	CAGAAACATATAGAACCTCCGGCAGAGATCAATGACGGCAGACCCAGTGGCCCGTG	778
Qy	287	GIYleuAePThrIleuAIIePrcISePrlIeAePrgIYsGIuAIIaValGIuGIuYsGIIn	306
Db	779	GGCTTAAACCCCTGGCTTCCCAAGCATCAACCGAAGGAAGTGTGGAAGAAAGAGC	838
Qy	307	ProTPraIAThYleuSePrgIYhIsmcAIIeGIYrThIsmcAIIeTPrgIYhIsmcAIIe	326
Db	839	CCCTGGGCAATCTGAGCTGGAGCATGTGACCGGATTCACACAGCTGGGGCCATCGAGC	898
Qy	327	AsPThrGIuAIIaGIuAIIePrgIYsPrcIYpheaPrgThrValGIYPrIYrValIPro	346
Db	899	GAGCGGAGGACGACGAGAGGAGTGTCCATGTGAGAGACTGTGGGCCCTTAGTCCCT	958
Qy	347	LeuTPrluGIYCYsGIuAIIaGIYpheaPrgThrValaPrlYpPrIYrThIsmcAIIePhe	366
Db	959	CTGTGGCTGGCTGTGAGCAGGATTTTATGTTCATGCGGGACCCCAACATCAGGCTTTC	1018
Qy	367	ThrPrcIYsGIYhIsmcAIIePrgIYsSePrlYsSePrlYrThIsmcAIIeTPrgIYhIsmcAIIe	386
Db	1019	ACCCCTGGGGGACGCTGTTCAGAAAGTCTGCAAGATCTGTGTGAGATCCCACTG	1078
Qy	387	ProIIeGIYrThIsmcAIIePheAIIaAIIeCYsPrcIYpheaPrgIYhIsmcAIIeValGIY	406
Db	1079	CCCAACGGAACGACGGGTTTCAGCGCGCTGTCTGTCTGGCAGCGAGCTGTGTGT	1138
Qy	407	GIuGIInAePrcIYleuYsIePheGIInGIYPrIYrThIsmcAIIePheAIIe	420
Db	1139	GAAACAGAACTGCATCAATGTATTTTCCAAAGTCTGACTGAC	1180

DEFINITION Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230208C14 product:pellino 2, full insert sequence.

ACCESSION AK045515

VERSION AK045515.1

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)

REFERENCE Carninci, P. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisake, S., Inoue, K., Togawa, Y., Ozawa, K., Ohara, E., Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multipillar sequencer Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

REFERENCE The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

REFERENCE Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details.

FEATURES

source

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

1..3530

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:B230208C14"

/db_xref="taxon:10090"

/clone="B230208C14"

/sex="male"

/tissue_type="corpora quadrigemina"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

292..1421

/note="pellino 2 (MGD|WGI:1891445, GB|NM_033602, evidence: BLASTN, 99%, match=864)

3513..3518

/note="putative"

3530

/note="putative"

3530

/note="putative"

ORIGIN

Alignment Scores:

Pred. No.: 1.23e-182 Length: 3530

Score: 1926.00 Matches: 361

Percent Similarity: 87.4% Conservative: 6

Best Local Similarity: 86.0% Mismatches: 9

Query Match: 84.1% Indels: 45

DB: 4 Gaps: 2

US-10-041-030-4 (1-420) x AK045515 (1-3530)

QY 1 MetPhaserProGlyngingluGluH1ScyAlaProAsnlySGluproVallySTyGly 20

DB 292 ATGTTTCCCGGGGCGAGAGGAAACCAGCGCCCAACAGAGAGCGGTGAAATACGGG 351

QY 21 GluLeuValValleuGlyTyraenGlyAlaLeuProAsnGlyAspArgGlyArgGly 40

DB 352 GAGCTGGTGGTCTT-GGG----- 368

QY 41 SerArgPheAlaLeuTyryAsArgProlyAlaAsnGlyVallySerProSerThrValHis 60

DB 368 ----- 368

QY 61 ValIleSerThrProGlnAlaSerIyAlaIleSerCyelySGlyGlnHisSerIleSer 80

DB 369 -----GCCATGAGCTCCAGAGGACATCCACATATCG 401

QY 81 TyrThrLeuSerArgAsnGlnThrValValValGlyTyThrHisAspLyAspThrAsp 100

DB 402 TACACGTTGCACGAGCAGACGAGTATGAGTGAAGTACACACATTAAGACACAGAC 461

QY 101 MetPheGlnValGlyArgSerThrGlySerProIleAspPheValValThrAspThrIle 120

DB 462 ATGTTTCAGGTGGCGAGGTTCACAGAAAGCCCATTTGATCTCTGTGTACACACGCGT 521

QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140

DB 522 TCCGGCGGTGAGAACGAAGT---GCCCATGATACCGACAGGACCATCTTACGTTCCGA 578

QY 141 CysArgIleValCysAspArgAsnGluProTyThrAlaArgIlePheAlaArgIlePhe 160

DB 579 TGCAGAGTCTGTGTGACAGGAACGACCATACAGACGCAATATTCGGCAGAGATTC 638

QY 161 AspSerSerIyAsnIlePheLeuGlyGluValAlaAlaIySTyPlyAsnProAspGly 180

DB 639 GATTCTTCCAAAATATCTTTCTTGGAGAGAAAGCAAAATGCAAAACCTGATGGA 698

QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200

DB 699 CACATGATGATCTCATTACCAACGAGTGTCTGTATGACACCCGCAAGAGGCTTCACC 758

Oy		201	GluGusSerInPProGlYValTPraPrGluIleSeSerValCYSGlYAspValTYThreU	220
Db		759	GAGCAATCCACGCCCTGGAGTCTCGAGAAGAGATCTGTCTGTGGGGAATGTATCACCTTG	818
Oy		221	ArgLutThrAgsSerAlaGlnGlnArgGlyLYLeuValGIsuSerGluThrAsnValIeu	240
Db		819	CGAGAGAACCAAGTCGGCCCCACAGAGGGGAAAGCTGTGGAAAGTGAACCAAGTCCTGT	878
Oy		241	GlnAspGlysSerIleuIlaAspLeuCysGlyValArthrLeuLeuTParXTThrAlaAspGly	260
Db		879	CNAGAGGCCTCCTCATTAAGACTGTGTGGGGCCACTTCTCTGTGAGATGCGAGATGGC	938
Oy		261	LeuPheHisThrProThrGlnLYLeuHsiTrIGluAlaLeuAArgGlnGluIleAsnAlaAla	280
Db		939	CTTTTTCACGCTCCTCACTCAGAGAACAATTAAAGCCCTCCGCGAGAGATCAAATCGACC	998
Oy		281	ArgProGlnCYsPProValAGlyLeuAsnThrLeuAlaPheProSerTleaAsnArgLYsGlu	300
Db		999	CGACCCCAGTGCCTCCGTGGGCTTTAAACCCCTGGGCTTCCACAGATCAACGGAAAGGA	1058
Oy		301	ValValGluGluLYuVgInProTPalalyrTleuSerCYsGlyHsiValHisGlyTYrHis	320
Db		1059	GTGGGGAGAGAGAACAGACCCCTGGGCAATPCTGAGCTCGGCACATGCGACGGCTAACAC	1118
Oy		321	AenTPrgLYHsiAgsSerAAspThrGluIlaAsnGluAArgGluCYsProMetCYsArgThr	340
Db		1119	AGCTGGGGCCCATCGGAGGACACACGGAACCCAAGAGAGAGTGTCCATGTGCAGACT	1178
Oy		341	ValGlyProTYrValProLeuTPreUGlyCYsGluAlaGlyPheTYrValAspAlaGly	360
Db		1179	GTGGGCCCCCTCAAGTCCCTCTCGGTGGGCTGTGAGGACGAGATTTTAATGTCAGTGGGGA	1238
Oy		361	ProProThrHsiAlaPheThrProCYsGlyHsiValCYsSerGluLYsSerAlaLYsTYr	380
Db		1239	CCCCCACTCAAGCTTTTCAACCCCTCGGGGACAGTCTGTTCAGAGAACTCTGCCAAGTAC	1298
Oy		381	TTPSerGlnIleProLeuProHsiGlyThrHsiAlaPheHsiAlaAlaCYsAProPheCYs	400
Db		1299	TGATGCGAGATCCCACTGCCACCGACCGAAGACGACCGGTTTTCAAGCGCGCTGTCGGTTGCG	1358
Oy		401	AlaThrGlnLeuValGlyGluGlnAsnCysIleLYsLeuIlePheGlnGlyProIleAsp	420
Db		1359	GCCACGCGACGCTGGTGTGGAAACAGAACTGCATCAATGATTTCCAAGGTCCACGTGCAC	1418
RESULT 5				
AK045673				
LOCUS		2564 bp	mRNA	linear HTC 03-APR-2004
DEFINITION	Mus musculus adult male corpora quadrigemina cDNA, RIKEN			
	full-length enriched library, clone:B23020E21 product:pelJino 1,			
	full-insert sequence.			
ACCESSION	AK045673.1	GI:26337532		
VERSION	AK045673.1			
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;			
	Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE				
AUTHORS	Carninci,P. and Hayashizaki,Y.			
TITLE	High-efficiency full-length cDNA cloning			
JOURNAL	Mech. Enzymol. 303, 19-44 (1999)			
PUBMED	10349636			
REFERENCE				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,			
	Itoh,M., Kono,H., Okazaki,Y., Watanabe,M. and Hayashizaki,Y.			
	Normalization and subtraction of cap-trapper-selected cDNAs to			
	prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
PUBMED	11042159			
REFERENCE				
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,			

TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHORS
4	THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE PANTOM CONSORTIUM.	FUNCTIONAL ANNOTATION OF A FULL-LENGTH MOUSE CDNA COLLECTION	Nature 409, 685-690 (2001)	
5	THE PANTOM CONSORTIUM AND THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE I & II TEAM.	ANALYSIS OF THE MOUSE TRANSCRIPTOME BASED ON FUNCTIONAL ANNOTATION OF 60,770 FULL-LENGTH CDNAS	Nature 420, 563-573 (2002)	
6	(BASES 1 TO 2584)			
7	ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P., FUKUDA, S., FUZURO, M., HANAGATA, T., HARA, A., HASHIZUME, W., HAYASHIDA, K., HAYASE, N., HIRAMOTO, K., HIRAKATA, T., HIROZANE, T., HORI, F., IMOTO, K., ICHI, Y., ITOH, M., KAGAWA, I., KASUKAWA, T., KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KOMO, H., KOSAKI, M., KOYA, S., KUNIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M., NAKAMURA, M., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHNATO, N., OKAZAKI, Y., SAITO, R., SAITOH, K., SAKAI, C., SAKAI, K., SAKAZUME, N., SANO, H., SAKAKI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T., SUGIYAMA, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU, K., TAKAHARA, S., TAKEDA, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A., YAMAMOTO, R., YAMAMOTO, H., SAKAGUCHI, S., IKEGAMI, T., KASHIYAGI, K., FUJII, K., INOUE, K., TOGAWA, Y., IZAWA, M., OHARA, E., WATABE, M., YONEDA, Y., ISHII, K., OZAWA, K., TANAKA, T., MATSUURA, S., KAWAI, J., OKAZAKI, Y., MURAMATSU, M., INOUE, Y., KITA, A. AND HAYASHIZAKI, Y.			
8	RIKEN INTEGRATED SEQUENCE ANALYSIS (RISA) SYSTEM-384-FORMAT	SEQUENCING PIPELINE WITH 384 MULTICAPILLARY SEQUENCER	Genome Res. 10 (11), 1757-1771 (2000)	
9	11076861			
10	11076861			
11	11076861			
12	11076861			
13	11076861			
14	11076861			
15	11076861			
16	11076861			
17	11076861			
18	11076861			
19	11076861			
20	11076861			
21	11076861			
22	11076861			
23	11076861			
24	11076861			
25	11076861			
26	11076861			
27	11076861			
28	11076861			
29	11076861			
30	11076861			
31	11076861			
32	11076861			
33	11076861			
34	11076861			
35	11076861			
36	11076861			
37	11076861			
38	11076861			
39	11076861			
40	11076861			
41	11076861			
42	11076861			
43	11076861			
44	11076861			
45	11076861			
46	11076861			
47	11076861			
48	11076861			
49	11076861			
50	11076861			
51	11076861			
52	11076861			
53	11076861			
54	11076861			
55	11076861			
56	11076861			
57	11076861			
58	11076861			
59	11076861			
60	11076861			
61	11076861			
62	11076861			
63	11076861			
64	11076861			
65	11076861			
66	11076861			
67	11076861			
68	11076861			
69	11076861			
70	11076861			
71	11076861			
72	11076861			
73	11076861			
74	11076861			
75	11076861			
76	11076861			
77	11076861			
78	11076861			
79	11076861			

ORIGIN

SKNIPLEGKAKWKTSDGMDGLTTNGVLVWHPRNGFTEDSKPGEIWEISVCGNVSFL
RFRSAOQKRGKWEIETNOLQDSLIDLCATLIMPARBELSTPRYKHEALROBIN
RARPQCPVGRNTLAFBMRKRVYVDEKQPVYVINCGRVETNMGWNEEDKQREBP
MCSVGPYVPLMLGCEAGFYVADGPPHASPCHGVCSEKTTAWSOIPLPHQTHTEH
AACPFCAHQIAGGQYIRLIFQGPLD"

Alignment Scores:

Pred. No.: 1.98e-181 Length: 2584
Score: 1912.00 Matches: 341
Percent Similarity: 90.5% Conservative: 39
Best Local Similarity: 81.2% Mismatches: 38
Query Match: 83.5% Indels: 2
Gaps: 1

US-10-041-030-4 (1-420) x AK045673 (1-2584)

Qy 1 MetPheSerProGlyInGluGluHicCysAlaProAsnYsgIuProValYsYrGly 20
Db ATGTTTCTCTCTATCAAGAAATCAT-----CTTCCAAAGCCCAATTAATATGCG 479
Qy 21 GluLeuValIleuGlyYrAsnGlyAlaLeuProAsnGlyAspArgIyArgArgLy 40
Db GAACCTCATGCTTGAATATGATGATCTCCCAACGGTATGAGGAGGAGGAAA 539
Qy 41 SerArgPheAlaLeuYrYlyAsnGlyProGlyAlaAsnGlyValYsProSerThrValHis 60
Db AGAAGGTTTCTCTTAAAGACCTTAAGGCAATGCGGTGAAGCCCTTACACCGTCCAC 599
Qy 61 ValIleSerThrProGluAlaSerYsAlaIleSerCysYsGlyInHicSerIleSer 80
Db ATTCGATGATCTCTCGAGCTGCGCAAGCAATTAAGCAACGACCAATGATATCA 659
Qy 81 TyrThrLeuSerArgAsnGluInHicValIleValGluYrThrHisAspYsAspThrAsp 100
Db TATATCTTATCTCGAGCCCAAGCGTGTGTTGATATCTCATGACAGACACTGAT 719
Qy 101 MetPheGluValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
Db ATGTTTCAGATGTCGTCGTCACGAAGGTCCTATTGATTTTGTAGTACGACACCGTT 779
Qy 121 SerGlySerGluInHicAsnGluAlaGluIleThrGluInSerThrIleSerArgPheAla 140
Db CTGGAAGTGAAGTAAATCCGACACGACGACGATTAAGCAAGCACTATATCAAGTTTGGC 839
Qy 141 CysArgIleValCysAspArgAsnGluProYrThrAlaArgIlePheAlaIleGlyPhe 160
Db TGTAGATCATATGTGAGCGGACGTCCTTTACGCTGGATTTATGCTGCGAGGTTT 899
Qy 161 AspSerSerYsAsnIlePheLeuGlyGluYsAlaIleYsYrPlyAsnProAspGly 180
Db GATTCATCAAAAACATCTTCTTGGGAGAGAGGCTCCCAAGTGAAGACATCGATCGG 959
Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db CAGATGATGCTGTCGACCACTAAAGAGTCTTGTATGATCACTCACTAATGCGTTTACA 1019
Qy 201 GluGluSerGluProGlyValIleThrArgGluIleSerValCysGlyAspValYrThrLeu 220
Db GAAAGCTCCAAACCTGGAATATGAGAAATATCAATATGCGAAAGCTTCTCAAGCTG 1079
Qy 221 ArgGluThrArgSerAlaGluInArgGlyYsLeuValGluSerGluThrAsnValLeu 240
Db CGTGAAACAGATCAGCTCAGCAGAGAGAAAGATGTTGAGAAATGAAACCAATCACCTTA 1139
Qy 241 GluAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
Db CAAAGTGGCTCTTAATGACCTTGGTGCAACTTGTGCGCTACTGCGAAGAGGC 1199
Qy 261 LeuPheHisThrProThrGluInYsHisIleGluAlaLeuArgGluInIleAsnAlaIle 280
Db CTTTCCATATCTCTACTGAGACACTTAAGACCTTTAAGACAGAGATCATATGACGCT 1259

Qy 281 ArgProGluCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLyGlu 300
Db CGGCGGAGTGGCTCTTATGAGGTTCAACACATGACCTTCCCGATTAAGAGAGAGAT 1319
Qy 301 ValIleGluGluYsGluInProIleAlaYrLeuSerCysGlyYsIleValHisGlyYrHis 320
Db GTTGTAGAGAAAGAACCAATGATATCTTAATCTGCGGCAATGCTCATGTTATCAT 1379
Qy 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGlyCysProSerCysArgThr 340
Db AACTGGGAAACAAAGAACGTCGCGCAAGATGCGAATGCTCATATGTTAGGCTCT 1439
Qy 341 ValGlyProYrValProLeuThrLeuGlyCysGluAlaGlyPheYrValAspAlaGly 360
Db GTTGTCCTATATCTCTGTCGCTTGAATGAGCTGATTTATGAGAGCGCGGC 1499
Qy 361 ProProThrHisAlaPheThrProCysGlyYsIleValCysSerGluYsSerAlaYsYr 380
Db CTTCCACCATGCTCTTATGAGCCCTGTCGCACTGTGTTCAGAAAGACAGCGGCTTAC 1559
Qy 381 TrpSerGluIleProLeuProHisGlyYrHisAlaPheHisAlaIleAspProPheCys 400
Db TGGTCCAGATCCCACTTCTCATGATGAGCAGCACTTTTCATGACAGCTGCGCTCTGT 1619
Qy 401 AlaThrGluLeuValGlyGluInAsnGlyIleYsLeuIlePheGluInGlyProIleAsp 420
Db GCACATCATGTCGTCGTGTAAGAGCTATATCAACTTATTTCCAGAGACCTTTAGAC 1679

RESULT 6

AY409117 1186 bp DNA linear GSS 16-DEC-2003
LOCUS
DEFINITION Pan troglodytes PELI2 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY409117.1 GI:39765085
GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pan.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

1 (bases 1 to 1186)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Fertler, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment
Location/Qualifiers
1..1186
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1186
/gene="PELI2"
/locus_tag="HCMA450"

ORIGIN

Alignment Scores:
Pred. No.: 1.3e-158 Length: 1186
Score: 1681.00 Matches: 317

Percent Similarity: 80.5%
 Best Local Similarity: 80.5%
 Query Match: 73.4%
 DB: 10
 Gaps: 0

US-10-041-030-4 (1-420) x AY409117 (1-1186)

QY 27 TyrAnG1yAlaLeuProAsnG1yAspArgG1yArGArGlySerArGpheAlaLeuTyr 46
 Db 2 TACAATGCTGCTTAACCAATGAGATAGAGACGAGAAATAAGATTGGCCCTTAC 61
 QY 47 LysArgProLysAlaAnG1yValLysProSerThrValHisVal11eSerThrProGln 66
 Db 62 AAGCGCCCAAGGCAAAATGGTGCAAAACCCGACATCTGCTGATGATCCACGCCCCAG 121
 QY 67 AlAserYsAla1eSerCyAlvG1yGlnHisSer1eSerYrThrLeuSerArgAsn 86
 Db 122 GCGTCCAAAGGCTACAGCTGCMAAGGTCACACAGTATCTTACACTTGTCAAGGAAT 181
 QY 87 GlnThrValVal1G1yTyrThrHisAspLysAspThrAspMetPheGlnVal1G1yArg 106
 Db 182 CAGACTGTGGTGGTGAAGTACACATGATAGAGACACGATATGTTCCAGNNNNNNNN 241
 QY 107 SerThrGluSerPro1LeaPheValVal1ThraPThr1eSerG1ySerGlnAsnThr 126
 Db 242 NNN 301
 QY 127 AspGlnAlaGln1eThrGlnSerThr1eSerArgPheAlaCyAspArg1eValCyAsp 146
 Db 302 NNN 361
 QY 147 ArgAsnGluPro1yThrAlaArg1ePheAla1aG1yPheAspSerSerLysAsn1e 166
 Db 362 NNN 421
 QY 167 PheLeuG1yG1yValAlaAlaLysTyrLysAsnProAspG1yHisMetAspG1yLeuThr 186
 Db 422 NNN 481
 QY 187 ThrAsnG1yValLeuValMetHisProArgG1yG1yLeuThrG1yGluG1ySerGlnProG1y 206
 Db 482 ACTAATGCGCTTCTGGTGAATGATCCACGAGGGGCTTCCACCGAGAGTCCACCGCCGG 541
 QY 207 ValTyrArgGlu1eSerValCyG1yAspValTyrThrLeuArgG1yThrArgSerAla 226
 Db 542 GTCTGGGGCGAGATCTGTCTGTGTGAGATGTGTACACCTTGCAGAAACAGGTCCGCC 601
 QY 227 GlnGlnArgG1yLysLeuValGluSerG1yThrAsnValLeuGlnAspG1ySerLeu1e 246
 Db 602 CAGCAAGAGAGAAAGCTGTGGAAGTGAGACCAACGCTCCGACGAGCGGCTCCCTCAT 661
 QY 247 AspLeuCyAspG1yAlaThrLeuLeuTyrArgThrAlaAspG1yLeuPheHisThrProThr 266
 Db 662 GACTGTGTGGGGCCACTCTCTCTGAGAAACAGCAATGGGCTTTTCACTCCCACT 721
 QY 267 GlnLysHis1eG1yAlaLeuArgG1yGln1eAsnAla1aArgProGlnCyAspProVal 286
 Db 722 CAGAGACATAGAACGCTCCGCGCAGAGATTAAACGCCGCCGCTCAGTGTCTGTG 781
 QY 287 G1yLeuAsnThrLeuAlaPheProSer1eAsnArgLysG1yVal1G1yGluG1yGln 306
 Db 782 GGGGCTCAACACCTGGGCTTCCCGACATCAACGAAAGAGGTGTGAGAGAGAAACAG 841
 QY 307 ProTyrAlaTyrLeuSerCyAspG1yHisValHisVal1eG1yThrHisAsnTyrG1yHisAspSer 326
 Db 842 CCCCTGGGCAATATCTCAATTTGGCCACGTGACCGGTATCCCAACTGGGGCCATCCGAGT 901
 QY 327 AspThrGln1aAsnG1yArgG1yGluCyAspProMetCyAspArgThrVal1G1yProValPro 346
 Db 902 GACACGAGGCGCAACGAGAGGAGTGTCCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 961
 QY 347 LeuTyrLeuG1yCyAspG1yAlaG1yPheTyrValAspAlaG1yProProThrHisAlaPhe 366

Db 962 CTCTGGCTGGCTGTGAGGAGGAGATTATGTAGACGACGAGACCGCACTNNNGCTTTC 1021
 QY 367 ThrProCyAspG1yHisValCySerSerG1yLysSerAlaLysTyrThrSerGln1eProLeu 386
 Db 1022 ACTCCCTGTGGAACGTGTCTCGAGAACTCTGCMAAATACTGTCTGATCCCGCTTG 1081
 QY 387 ProHisG1yThrHisAlaPheHisAlaAlaCyAspProPheCyAsnAlaThrGlnLeuVal1G1y 406
 Db 1082 CTCATGAAACTCATGATTCACGCTGCTGGCTTCTGTCTGCTACACAGCTGTTGGG 1141
 QY 407 GlnGlnAsnCyAsp1eLysLeu1ePheGlnG1yPro1LeaP 420
 Db 1142 GAGCAAAACTGCATCAATTAATTTTCAAGGTCAATTGAC 1183

RESULT 7
 AK044418
 LOCUS
 DEFINITION
 AK044418 1879 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult retina cDNA, RIKEN full-length enriched library,
 clone:A930011L17 product:similar to PBLINO 2 (PBLINO (DROSOPHILA)
 HOMOLOG 2) (Homo sapiens), full insert sequence.
 AK044418
 VERSION
 AK044418.1 GI:26336455
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 10349636

REFERENCE
 AUTHORS
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 11042159

JOURNAL
 PUBMED
 11042159

TITLE
 SHIBATA, K., ITOH, M., AIZAWA, K., NAGAOKA, S., SASEKI, N., CARNINCI, P.,
 KOMNO, H., AKIYAMA, J., NISHI, K., KITSUNAI, T., TASHIRO, H., ITOH, M.,
 SUMI, N., IEBHI, Y., NAKAMURA, S., HASEMOTO, M., NISHINE, T., HARADA, A.,
 YAMAMOTO, R., MATSUMOTO, H., SAKAGUCHI, S., IKEYAMA, T., KASHIWAGI, K.,
 FUJIIAKE, S., INOUE, K., TOGAWA, Y., IZAWA, M., OHARA, E., MATSUURA, S., KAWAI, J.,
 YONEDA, Y., ISHIKAWA, T., OZAWA, K., TANAKA, T., MATSUURA, S., KAWAI, J.,
 OKAZAKI, Y., MURAMATSU, M., INOUE, Y., KITA, A. and HAYASHIZAKI, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multiplexed sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 11076861

JOURNAL
 PUBMED
 11076861

TITLE
 THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE
 FANTOM CONSORTIUM.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

JOURNAL
 PUBMED
 11076861

TITLE
 THE FANTOM CONSORTIUM AND THE RIKEN GENOME EXPLORATION RESEARCH
 GROUP PHASE I & II TEAM.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

JOURNAL
 PUBMED
 11076861

TITLE
 ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P.,
 FUKUDA, S., FURUKO, M., HANAGAKI, T., HARA, A., HASHIZUME, W.,
 HAYASHIDA, K., HAYATSU, N., HIRAMOTO, K., HIRAOKA, T., HIRAZANE, T.,
 HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T.,
 KATOH, H., KAWAI, J., KOJIMA, Y., KONDO, S., KOMNO, H., KOUDA, M.,
 KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M.,
 NAKAMURA, M., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHSAO, N.,
 OKAZAKI, Y., SAITO, R., SAITO, H., SAKAI, C., SAKAI, K., SAKAZUME, N.,
 SANO, H., SASEKI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T.,

TITLE

Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S.,
Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A.,
Muramatsu M., and Hayashizaki Y.

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Retina RNA was provided by Dr. Stefano Gustincich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA
02115, USA) whose assistance is gratefully acknowledged. Please
visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

FEATURES

source

CDS

```
1. 1879
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A930011L17"
/db_xref="taxon:10090"
/clone="A930011L17"
/tissue_type="retina"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
104..1441
/note="unnamed protein product; putative
similar to PELLINO 2 (PELLINO (DROSOPHILA) HOMOLOG 2)
[Homo sapiens] (SPTK1/SHATB, evidence: FASTA, 71.1%ID,
99.2%length, match=1245)"
/codon_start=1
/protein_id="BAC1913.1"
/db_xref="GI:26336456"
/translation="MTLEKNDVDSPTSDLDHSGSGCGLSGPGEALAGEPIKY
GELIVAGNCLASGDKRRSRSLASRRHANGYKRDWENHISTEYVSLNRGCH
SISPTLSRSHSVIEYTHSDKDFQIGSTENNIDVVDTSFGGATGSPASQST
SYACRILCDRREPPTARIYAAGFDASNTFLERAKAKMTPLGLMDGLIDLVNH
PAGGFSDSAPGWRBSVCGNYTLEDSRSAOQRGLVNESNVLDGSLIDCGAT
LWRTPLGLRAPTLKOLEAORORANARPOCPVGTSLAPSPARGRTPDQOPW
YVRCGHVGHGWCRCREOGRPOREBCPLGVGVYVLMIGOEAGCLDGPSSHAFV
PCGHVSEKTRARYAOTPLPHGTAFHFAACPFCAWLTGELGCVRLIFGSLPDL"
```

ORIGIN

Alignment Scores:

Pred. No.: 2,34e-153 Length: 1879
Score: 1632.00 Matches: 297
Percent Similarity: 81.6% Conservative: 44
Best Local Similarity: 71.1% Mismatches: 73
Query Match: 71.3% Indels: 4
DB: 4 Gaps: 3

US-10-041-030-4 (1-420) x AK044418 (1-1879)

```
QY 4 Progl1yngingluglunh1scyalaprobasmlysglubrovallytytygylglu1eVal 23
Db 194 CCGGTGAAGA-----GCACCTGGCAGGCGAGGAGCCATCAAGTATGGTGAATCATC 247
QY 24 Val1euglytytyanglyalaleuprobasmlyseparglyalaglysestaryphe 43
Db 248 GTTCTGGGCTACATGGGTCTCTGGCAAGTGGACAGAGGCCGCCCGAGAGCCGCTG 307
QY 44 AlaleutyrylsaypPOLysalalaanglyValylsPProserThrvallH1sval11Ser 63
Db 308 GCACCTAGGCGCGCCGACATGCGAGAGTGAAGCCAGATGTCATGACCAATCTCC 367
```

```
QY 64 ThProgl1alaserlysal1aleserCylyeglygl1h1ser1leserTytrleu 83
Db 368 ACACCACTGCTTCACAGGCCCTGAGTACCGAGGCCAGACATCTTATTCACACTG 427
QY 84 SerArgasnglnThrvallValValgluTytrThH1sAspLyasPThraPmetPheGln 103
Db 428 TCCCGAGTCACTACATGATGAGTACACATGACAGGACAGCAAGCATGTTCCAG 487
QY 104 ValGlyArgserThrgluserPro1leasPheValValThrasPThrlleserGlyser 123
Db 488 ATTGGCCGCTTACTAATAAAGATTTGTTGTGAACACATCTCTGCTGGA 547
QY 124 GlAsnThrasPgl1alagln1leThrglnserThrlSeraPhealCyasGly1e 143
Db 548 GGGGCTACA---GAGGCGCTTGTGCCAAGTACATCTCCGCTATGCTTCAGAAAT 604
QY 144 ValCyasPArgasnglnupProTytrThralArg1lePhealalaglyPheasPser 163
Db 605 CTCTGTGACCGCGGACACCTTATCTGCCCGAATCTAGCGCTGCTTGTATGCTCT 664
QY 164 LysAsn1lePheleuglyglulysalalalytyrlyshasPProasPglyH1sMetAsp 183
Db 665 AGCAACATCTTCTTGAGAGCGCGGACCAATGAGCACTCCGATGCTGATGAG 724
QY 184 GlyLeuThrvanGlyVal1eValMetH1sProargGlygllyPheThrglugluser 203
Db 725 GGCTTGAACAACATGGGCTTGTGATGACCCAGAGTGGCTTCTCGAGACTCT 784
QY 204 GlPProgl1Val1trParglul1eserValCyglYasPlyTytrThleuArgluThr 223
Db 785 GCCCGAGTCTCGAGAGGATTTCTGTGCGGAAATGTACACGCTCGCGGATAGC 844
QY 224 ArgSer1agln1narglylyseuValgluserGluThrasVal1euglnasPgly 243
Db 845 CGATCAGCTCAGAGCGGGGAGCTGTGAGAAAGAAATCCAACTCTCGCAAGCGGT 904
QY 244 Serleu1leasPleuCyglYal1atHrleu1eThrvanGlyleuPheH1s 263
Db 905 TCCCTACATCAGCTGTGTGGGACACATGTTGGCGACTCCGCTGGTGTCTGAGG 964
QY 264 ThProThrgln1nysH1s1leGlu1aleuArgGln1ul1easna1alAatPProGln 283
Db 965 GCACCACTGAAACAATGAGGCCAGGACAGGACAGGACAGGACAGGACAGGACAG 1024
QY 284 CyasPProvalGlyleuansThrleual1aPheProser1leasnarglysgluValGlu 303
Db 1025 TGGCCGCTGGGCTTACACCTTGGCTTCCAGGCCAGGCCGCTGCGCAGACAGCACC 1084
QY 304 Glulys---GlnProThralatTytrleuserCyglYH1sVal1H1sGlyTytrH1sAsnTTP 322
Db 1085 GACAAACAGCAACCTGGGCTACGTCCTTGTGTCATGTCATGATGCTATGACAGGCTGG 1144
QY 323 GlyH1sArgserasPThrglu1alaangluArgGlyCyasPmetCyasArgThrvallGly 342
Db 1145 GGCTGCCGAGGAGAACAGGCCCGCCAGAGCGAGAGTCTCTTCTGCGCCTTGTGGA 1204
QY 343 ProTytrValProleutryleuglyCysglu1aglyPheTytrValasParglyProPro 362
Db 1205 CCTATGTGCCCTGTGGCTCGGTGAGAGCGCGGTCTGCTGAGCCCTGAGCCACCC 1264
QY 363 ThrH1sAlaPheThrvProCyglYH1sValCyasPserGlyulysSer1alalyTytrTPser 382
Db 1265 AGCCACGCTTTTACACCTGTGACAGCTGTCTGTGAGAGACCTGCCCTACTGGGCT 1324
QY 383 Gln1leProleuProH1sGlyThrvH1sAlaPheH1sAla1a1aCyasPProPheCyasAlaThr 402
Db 1325 CAGACACGCTGCGGACAGGACCAACGCTTTCAGAGCTGCTGCTCTTCTGTGAAGCT 1384
QY 403 GlnleuValGlyglu1nansCy1leYleu1lePheGln1yPro1leasP 420
Db 1385 TGGCTACCGGTAGAGTGTGCTGTGCTGCTGCTATTTTCAAGGGGACAGTGAAC 1438
```

RESULT 8

gene tries
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1242)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Clevello, D.R., Lu, F., Murphy, B., Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sinsky, J.J., Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
 FEATURES
 SOURCE 1. 1242
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>1242
 /locus_tag="HCM3341"
 gene
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,43e-152 Length: 1242
 Score: 1621.50 Matches: 297
 Percent Similarity: 80.9% Conservative: 42
 Best Local Similarity: 70.9% Mismatches: 73
 Query Match: 70.8% Indels: 7
 DB: 10 Gaps: 4
 US-10-041-030-4 (1-420) x AY408789 (1-1242)
 QY 3 SerProGlyGlnGluGlnHisCysAlaProAsnIsgluProValIsglyTyrGlyGluLeu 22
 Db 1 TCTCCCGCT-----GAGATGCCAGCCAGCCAGCCAGCCATCAATGATGAACTC 54
 QY 23 ValValIleuGlyTyrAsnGlyAlaLeuProAsnIsglyAspArgIYATGATgIysSerArg 42
 Db 55 ATGCTCTGGGGCTCAATATGTTGCTGGCAAGTGGGGCAAGGGCCCGCGGCAAGCGC 114
 QY 43 PheAlaLeuTyrIlyAsnGlyProIyAlaAsnGlyValIlyProSerThrValHisValIle 62
 Db 115 CTGGCACTGAGCCCGCGGTGCGACAGCCAGCGGTGAAGCCAGCTCATGTCACACATC 174
 QY 63 SerThrProGlnAlaSerIyAlaIleSerCysIsglyGlnHisSerIleSerIyThr 82
 Db 175 TCCACGCCG-----CTGCCACTGATACCTGCTCAGCAGCATCTTCTATACA 225
 QY 83 LeuSerArgAsnGlnThrValAlaValGluTyrThrIAspIyAspThrAspMetPhe 102
 Db 226 CTGTCCCGGAGCCACTCGGTCAATGTAATACATGATATACATGATACAGACATGTTT 285
 QY 103 GlnValGlyArgSerThrIsglySerProIleAspPheValIlyThrAspThrIleSerGly 122
 Db 286 CAGATTGGCCGCTCAGAGAGACATATGATCTTGTGTACAGACACCTCCCTGGA 345
 QY 123 SerIleAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArg 142
 Db 346 GGA---GGGGCTGCCAGGGCCCTTTCGCCAGAGCAACATCTCCCTATGCTTGGCC 402
 QY 143 IleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIsglyPheAspSer 162
 Db 403 ATCTCTGTGACCCCGCCGACCTTATATCTCCCGCATATATGCCGCTTGTGATGCC 462
 QY 163 SerIyAsnIlePheLeuGlyGlnIlyAlaIlyAspIyIyAsnProAspGlyIyHisMet 182
 Db 463 TCTTGCACAACTCTTCTTGGAGAGCGAGCCGCAATGGCGAGCCCAAGTGGCTGATG 522
 QY 183 AspGlyLeuThrThrAsnGlyValLeuValMetHisProAspGlyIyGlyPheThrGlnGlu 202
 Db 523 GATGACTGACACCAATGAGTCTGTATGACCCCGGAGCGGCTTCTTCCAGAGAC 582
 QY 203 SerGlnProGlyValIlyPArgGluIleSerValCysGlyAspValIyTyrThrLeuArgGlu 222

Db 583 TCAGCCCGCGGTCTCGCCGAGATCTCGTCTGTGGATGTGATCAATGCGCGAC 642
 QY 223 ThrArgSerAlaGlnGlnArgGlyIyIyLeuValGluSerGluThrAsnValIleGlnAsp 242
 Db 643 AGCCGCTACCCAGAGCGGGCGAGCTGTGTAAACAGATCCATCGTCTCGAGAC 702
 QY 243 GlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrPArgThrAlaAspGlyLeuPhe 262
 Db 703 GGTCTCTATGACCTGTGTGGGCGACACTGCTGTGGGCGACACCGCGGGCTGTG 762
 QY 263 HisThrProThrGlnIyHisIleGluAlaLeuArgGlnIyIleAsnAlaIlyArgPro 282
 Db 763 CGGAGTCCCACTGTAAGCAACTGAGAGCCAGCGCAGAGGCAATGACAGCGCGCC 822
 QY 283 GlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgIyGluValIle 302
 Db 823 CAGTGGCCGCTGGGCTCAGCACTGTGGCTTCCAGGCCAGCCCGTGGCCGACAGCG 882
 QY 303 GluGluIyS---GlnProTyrAlaTyrLeuSerCysGlyIyHisValHisGlyTyrHisAsn 321
 Db 883 CCGGACAAACAGAGCCCTGGTCTTACGTCCGCTGGGAGACCTCATGCTACACGCG 942
 QY 322 TyrGlyHisArgSerAspThrGlnAlaAsnGluArgGluCysProMetCysArgThrVal 341
 Db 943 TGGGGCTGCGCGGAGCGGGGCCCCAGAGCGGAGATGCTCTGTCGCCCTTGTG 1002
 QY 342 GlyProTyrValProLeuTyrIyGlyGlyIyGlyPheTyrValAspAlaIyPro 361
 Db 1003 GGGCTTATGTGCTCTATGCTTGGCAGAGGCGGCTGCTGAGCCCTGGAGCC 1062
 QY 362 ProThrIleAlaPheThrProCysGlyIyHisValCysSerGluIySerAlaIyTyr 381
 Db 1063 CTAGCATGCTTGTGACCTTGGCGGACAGCTGTCTGTAGAAAGATGCGCGCTACTG 1122
 QY 382 SerGlnIleProLeuProHisGlyIyThrHisAlaPheHisAlaIyCysProPheCysAla 401
 Db 1123 GCCAGACACACATGCGCCAGCCAGCCAGCCATGCTTCATGCCCTGCGCTTTCGCG 1182
 QY 402 ThrGlnLeuValGlyGlnGlnAsnGlyIleIyLeuIlePheGlnIyProIleAsp 420
 Db 1183 GCCTGCTTACCGGCGAGCATGGCTGCTGCTCATTTTCCAGGCGCGCTGAGT 1239
 RESULT 11
 HSM805479 2710 bp mRNA linear HTC 22-SEP-2004
 LOCUS HSM805479
 DEFINITION Homo sapiens mRNA, cDNA DKFp761P1024 (from clone DKFp761P1024).
 ACCESSION AL834395
 VERSION AL834395.1 GI:21740082
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2710)
 AUTHORS Bloembergen, H., Becher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C., Oanger, A., Fobo, G., Han, M. and Wiemann, S.
 CONSRM The German cDNA Consortium
 TITLE Direct Submission
 JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
 COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de, sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFp761P1024) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
 Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/c1.cgi?cloneID=DKFp761P1024
 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES

SOURCE

```

Location/Qualifiers
1.. 2710
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="RZPD:DKFZp761P1024"
    /db_xref="taxon:9606"
    /clone="DKFZp761P1024"
    /tissue_type="amygdala"
    /clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
    /dev_stage="adult"
    /note="pellino protein homolog 3"
1.. 2710
    /gene="DKFZp761P1024"
    .112.. .1521
        /gene="DKFZp761P1024"
        /codon_start=1
        /product="hypothetical protein"
        /protein_id="CAD39057.1"
        /db_xref="GI:21740083"
        /db_xref="UniProt/Swiss-Prot:O8N2H9"
        /translation="MWLEKGPVSGSPRTSLDRHGKSCVLVSFGDEAGQGEPIKYV  

        VELIVAGCEGEGEETAGCGYVTGRARSCNGLCLASDCKRRKRRLALRRRHANG  

        VAPDMHIISTPLVSKLSRSQGHISITYLRSHSVITATIDSDTFQIGRSTENW  

        IDPVAFTSPGGGAEEPSAKGSOSTISRACILCDRRPYTAIIYAAGDASSNTFLDE  

        RAARKRTEDGMGLDTNGVLVNHMPAGSEBSAPVMELISEVCNVATTDSADSAOO  

        RKGIVENSNNLOGSLIDLGCATLLMRPAGLLRAPTLGLEAORGANARPCCVV  

        GLSTLPAGEPALCGRPDPSPHQKQPVYVRCCHYGHWGCRRERGQRERCPLCUNRP  

        VYPLMLGQAEALCGRDPSPSHAFACGVCSSEKTARYMAQTPLPHGTAFPAACPCCG  

        MWLEGGECGCVALLIQGGLD"

```

ORIGIN

Alignment Scores:

Pred. No.:	6,66e-152	Length:	2710
Score:	1620.00	Matches:	299
Percent Similarity:	77.0%	Conservative:	42
Best Local Similarity:	67.5%	Mismatches:	74
Query Match:	70.7%	Indels:	28
DB:	4	Gaps:	4

US-10-041-030-4 (1-420) x HSM805479 (1-2710)

Qy	3	SerProGlyIgLingluGluHLeCysAlaProAbnlysgIuPProVallyeTyGlyGluLeu	22
Db	199	TCGCCGGT-----GAAGATGCCAGCCAGAGAGGCCATCAAGTAGGTAATC	252
Qy	23	ValValIleuGly-----	26
Db	253	ATGCTCTGGGATCTGTGAGGAAGAGTAGGAACCGAGGCTCAGAGAGGGAAAGTG	312
Qy	27	-----TyrAnGlyAlaLeuProAbnGlyAspArgGlyArg	38
Db	313	ACTGGCCCAAGGGCACAGCTGTGCTACATGTTGTCTGGAGAGGGGAGCAAGGGCGCG	372
Qy	39	ArgIysSerArgPheAlaLeuTyTyrLysArgProLysAlaIleuGlyValIysProSerThr	58
Db	373	CGGCGAAGCCGCTCGGCACTGAGCCGCGGATGCCAGCCAGCGGGGTAAACCCAGAGCTC	432
Qy	59	ValHisValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSer	78
Db	433	ATGACCAACATCTCCACGGCCGCTGCTCCCAAGGCACAGATTAACCCGGTCAACACAGC	492
Qy	79	IleSerTyThrIleuSerArgAsnGlnThrValValValGluTyThrHisAspLysAsp	98
Db	493	ATCTCGATACACTGTCCCGGAGGCACCTGGTCAATAGGAGATTAACACATGATAGCCAC	552
Qy	99	ThrAspMetPheGlnValGlyArgSerThrGlnSerProIleAspPheValIleThrAsp	118
Db	553	ACAGACATGTTCCAGATGATGGCCGCTCCACAGAAACATGATTGACTTGTGTGTAACAGAC	612
Qy	119	ThrIleSerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArg	138
Db	613	ACGTCCTCCGTAGAG--GGGCTGGCCAGAGGCCCTTCTGCCCAGAGACATATCTCCGC	669

QY	139	ThellaCysaAgglIeValCysAspArgAsnGluProGlyThrAlaAgglIePheAla	158
Db	670	TATGCTGGCGGCAATCTCTGTGACCGCGGCGACCCCTATCTGCGGCACTATGCGCGCT	729
QY	159	GIyPheAspSerSerLysAsnIlePheLeuGIyGluLysAlaAlaLysTrpLysAsnPro	178
Db	730	GGCTTCGATGCTCTAGCAACATCTTCTCTTGAGAGAGGAGCGGCAATAGCGGACCCCA	789
QY	179	AspGIyAlaIleMetAspGIyLeuThrThrAsnGIyValLeuValIleMetHisProArgGIyGIy	198
Db	790	GATGGCGCTGATGTAGTGACTGACCAACCAATGTGACTCTGTATGTGCACCGCGGCGGGCGC	849
QY	199	PheThrGIyGluIleSerGIyAsnProGIyValIThrArgLuuIleSerValCysGIyAspValTyr	218
Db	850	TTCTCCAGAGACTAGCGCCCGGGGTGTGTGGGGGAGATCTCGGCTGTGTGGAAATGTATAC	909
QY	219	ThrLeuArgGluThrArgSerAlaGlnGIyAlaArgLysLeuValGluSerGIyThrAsn	238
Db	910	ACATTTGCGGACACAGCCGCTCAGCCCAAGCGGGGCAAGCTGGTGGAAACAGTCCAC	969
QY	239	ValLeuGlnAspGIySerLeuIleAspLeuCysGIyAlaIThrLeuLeuTrpArgThrAla	258
Db	970	GTGCTGACAGAGCGGCTCTTCATGACCTGTGTGTGGGCGACACTGCTGTGGCGCACCG	1029
QY	259	AspGIyLeuPheHisIThrProThrGlnLysHisIleGluValAlaMetArgGlnGluIleAsn	278
Db	1030	GCGGGGGCTGTGCGGGGCTCCCACTAGAACACTGGAGGGCCCAAGCGGCAAGGCAAT	1089
QY	279	AlaAlaIleArgProGlnGlyCysProValGIyLeuAsnIThrLeuAlaPheProSerAlaAsnArg	298
Db	1090	GCAGCGGCGGCCCCAGTGCGCCGTGGGCGCTCAGCACTGTGGCTTCCCAAGCCAGCCCGT	1149
QY	299	LysGluValValGIyGluLys---GlnProTrpAlaTyrLeuSerCysGIyHisValHis	317
Db	1150	GGCGGCGACAGCGCCCAACAAACAGCAACCTCGGTCTACGTCCGCTCGGGCAAGTCCAT	1209
QY	318	GIyTyrHisAsnITrGIyHisAsnSerAspIThrGluAlaAsnGluArgGluCysProMet	337
Db	1210	GGTACACACGGCTGGGGCTGCGCGGAGACGGGGGCCCCAGGAGCCGCAATGCTCTTC	1269
QY	338	CysArgIThrValGIyProTyrIThrAlaProLeuITrLeuGIyCysGluAlaGlyPheTyrVal	357
Db	1270	TGCGGCTTGTGGGGCTTATGTGCTCTTATGTGCTTGTGGCAGAGGGCGGCTGTGCTG	1329
QY	358	AspAlaGlyProProIThrHisAlaPheIThrProCysGIyHisValCysSerGIyLysSer	377
Db	1330	GACCTGGGCGGCTTACGACATGCTTTGGACCTTGGCGGCAACGTCTCTCTGAAACACT	1389
QY	378	AlaLysTyrITrSerGlnIleProLeuProHisGIyIThrHisAlaPheHisAlaAlaCys	397
Db	1390	GGCGGCTACAGGGCGGACACCACTGCCCCACGGGACCCCATGTTCCATGGCGGCTGC	1449
QY	398	ProPheCysAlaIThrGlnLeuValGIyGluGlnAsnCysIleLysLeuIlePheGlnGly	417
Db	1450	CCCTTTTGGCGGGCGCTTACCGGGAGACATGCTGTCGTCGCTCATTTTCCACGAGGC	1509
QY	418	ProIleAsp 420	
Db	1510	CCGCTGGAT 1518	

RESULT 12

LOCUS AY408791 1242 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM3341 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY408791
VERSION AY408791.1 GI:39764759
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

TITLE	JOURNAL	REFERENCE	AUTHORS
Genome Res. 10 (10), 1617-1630 (2000)			
11042159			
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komoto, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.			
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed cDNA library			
Genome Res. 10 (11), 1757-1771 (2000)			
11076861			
4			
THE RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.			
Functional annotation of a full-length mouse cDNA collection			
Nature 409, 685-690 (2001)			
5			
THE PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs			
Nature 420, 563-573 (2002)			
6 (bases 1 to 2474)			
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hasegawa, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imomani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sobabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku, A., Kihira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.			
Direct Submission			
Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Physical and Chemical Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]			
COMMENT			
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.			
Please visit our web site for further details.			
URL: http://genome.gsc.riken.jp/			
URL: http://lancm.gsc.riken.jp/			
Location/Qualifiers			
1. 2474			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="CS7BL/6J"			
/db_xref="PANTOM_DB:9230114E15"			
/db_xref="taxon:10090"			
/clone="9230114E15"			
/sex="male"			
/tissue_type="epididymis"			
/clone_lib="RIKEN full-length enriched mouse cDNA library"			
/dev_stage="adult"			
<1. 854			
/note="unnamed protein product; pe11ino 2			
(MGDI GJ:1891445, GB NW_033602, evidence: BLASTN, 99%, match=864)			
putative"			
/codon_start=3			
/protein_id="BAC28485.1"			
FEATURES			
source			
CDS			

[illegible]

Db 723 GCCAAGTACTGTCGAGATCCACCTGCCACGAGACGACGCGTTTACGCCGCTGT 782
 Qy 398 ProPheCysAlaThrGlnLeuValGlyGlnAsnCysIleuLeuIlePheGlnGly 417
 Db 783 CCGTTCGCCCGCCACGACGCTGTTGGTGAACAGAACTGCATCAATTGATTTTCCAAAGT 842
 Qy 418 ProIleAsp 420
 Db 843 CCAGTGGAC 851
 RESULT 14
 LOCUS BX388547 849 bp mRNA linear EST 28-APR-2004
 DEFINITION BX388547 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 cDNA clone CS0DD006YF18 5-PRIME, mRNA sequence.
 ACCESSION BX388547
 VERSION BX388547.1 GI:30460950
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.
 1. (bases 1 to 849)
 Ii,W.B., Gruber,C., Jesse,J. and Polayer,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 7349.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0NT008Z06_T0713_1&c=7349.f.
 Location/Qualifiers
 source
 1. 849
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD006YF18"
 /cclone_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.8e-142 Length: 849
 Score: 1519.00 Matches: 272
 Percent Similarity: 98.6% Conservative: 1
 Best Local Similarity: 98.2% Mismatches: 4
 Query Match: 66.3% Indels: 0
 Gaps: 0
 US-10-041-030-4 (1-420) x BX388547 (1-849)
 Qy 140 AlaCysArgIleValCysAspArgAsnGluProGlyThrAlaArgIlePheAlaIleGly 159
 Db 1 GCCTCGACGATTCGTCGACAGAAAGAACCTTACACACACGATATTCGCGCGGGA 60
 Qy 160 PheAspSerSerLeuAsnIlePheLeuGlyGluValAlaIleValTyrPheAsnProAsp 179
 Db 61 TTGGACCTCTCCAAAACATATTTCTTGGAGAAAAGCGAAGAGTGAAGAAACCCGAC 120
 Qy 180 GlyHisMetAspGlyLeuThrThrAsnGlyValIleuValMetHisProArgGlyPhe 199

Db 121 GCCACATGAGTGGGCTCACTAAATGGCTCTGTGATGATCATCAGAGGGGCTTC 180
 Qy 200 ThrGluSerGlnProGlyValTyrArgGluIleSerValCysGlyAspValTyrThr 219
 Db 181 ACCGAGAGTCCCAAGCCCGGGGTCTGGCCGAGATCTGTCTGTGAGATGTGTACCC 240
 Qy 220 LeuArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnVal 239
 Db 241 TTCGAGAAACCAAGTCTGGCCGACGACGAGAAAGCTGGTGAAGTGAAGACCAAGTC 300
 Qy 240 LeuGlnAspGlySerLeuIleAspLeuCysGlyValThrLeuLeuThrArgThrAlaAsp 259
 Db 301 CTGCAGAGACGGCTCCCTCATATGACTGTGTGGGGCCACTCTCTCTGAGAAACAGAGAT 360
 Qy 260 GlyLeuPheHisThrProThrGlnValHisIleGluValAlaLeuArgGlnGluIleAsnAla 279
 Db 361 GGGCTTTTCAATCTCCAACTCGAAGGACATGAAAGCCCTCCGGCAGAGATTAACGCC 420
 Qy 280 AlaArgProGlnCysArgProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGly 299
 Db 421 GCCCGGCTCAGTCTCTGTGGGGCTCAACACCTGGCTTCCCGACATCAACAGGAA 480
 Qy 300 GluValValGluGluValGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyr 319
 Db 481 GAGGTGTGAGGAGAGAACGACCCCTGGCATATCTCAGTTGTGCCACGTGCACGGGTAC 540
 Qy 320 HisAsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArg 339
 Db 541 CACAACTGGGGCCATGAGTACAGACAGAGCCCAAGAGGAGAGTCCATGTGAG 600
 Qy 340 ThrValGlyProTyrValProLeuThrPleuGlyCysGluValAlaGlyPheTyrValAspAla 359
 Db 601 ACGTGGGGCCCATATGAGTCTCTGTGGCTTGTGTGAGGACGAGATTTATATGACGCA 660
 Qy 360 GlyProProThrHisAlaPheThrProCysGlyHisValCysSerGluValSerAlaVal 379
 Db 661 GGACCGCCAACTCATCTCTTCACTCCCTGTGACAGCTGTGGACAGAAAGTTCGAAA 720
 Qy 380 TyrTrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIleAspCysProPhe 399
 Db 721 TACTGGCTCAGATCCCGTGTGCTCATAGGAATCATGACATTTACCGCTGTTCCCTTTC 780
 Qy 400 CysAlaThrGlnLeuValGlyGlnAsnCysIleLeuLeuIlePheGln 416
 Db 781 TGCGTCAACACAGCTGGNTGGGAGCAAACTGCATCAATTAATTTTCAA 831
 RESULT 15
 CX975821/c 855 bp mRNA linear EST 08-FEB-2005
 LOCUS UGI CAAP9337.fwd NIH XGC tropint1 Xenopus tropicalis cDNA clone
 DEFINITION IMAGE:7715625 5', mRNA sequence.
 ACCESSION CX975821 GI:58789343
 VERSION CX975821.1
 SOURCE Xenopus tropicalis (western clawed frog)
 KEYWORDS Xenopus tropicalis
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 855)
 Richardson,P., Lucas,S., Rohnsar,D., Deter,J.C., Ng,D.C.,
 Brokstein,P. and Lindquist,E.A.
 DOE Joint Genome Institute Xenopus tropicalis EST project
 Unpublished (2004)
 TITLE JOURNAL
 COMMENT Other ESTs: UGI CAAP9337.rev
 Contact: Lindquist,E.A., Richardson,P.
 DOE Joint Genome Institute
 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 Tel: 925 296 5600
 Fax: 925 296 5710
 Email: cdna@jgi-psf.org
 Tissue Procurement: Robert M. Grainger
 cDNA Library Preparation: Bruce Blumberg Laboratory, University of

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 2, 2006, 04:45:41 ; Search time 809 Seconds

(without alignments)
3460.038 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MESPGEHRAHPKPEPVKVG.....ATQVGRQNCIKILPQPID 420

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xjh
-O=/abs/ABSSWEB.spool/US10041030/runat_01032006_134403_21809/app.query.fasta-1
-DB=N Geneseg -OFFT=fastap -SUPFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCNALGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=exc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs808
-USRR=US10041030 -CCGN 1.1 727 @runat_01032006_134403_21809 -NCPU=6 -ICPU=3
-NO MMP -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_21: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *
14: geneseqn2005s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2290	100.0	1565	6 AAL46323	AAL46323 Human M33
2	2290	100.0	1823	10 ADC30209	ADC30209 Human nov
3	2290	100.0	4563	6 ABC78319	ABC78319 Nucleotid
4	2290	100.0	5579	8 ACC42349	ACC42349 Human MAP

5	2283	99.7	1263	6 ABL58449	ABL58449 Human pel
6	2283	99.7	1263	12 ADP48670	ADP48670 Human pel
7	2175.5	95.0	1260	6 ABL58448	ABL58448 Murine pe
8	2175.5	95.0	1260	12 ADP48668	ADP48668 Mouse pel
9	2111.5	92.2	1717	8 ACC42348	ACC42348 Mouse MAP
10	1917	83.7	1257	6 ABL58447	ABL58447 Human pel
11	1917	83.7	1257	12 ADP48666	ADP48666 Human pel
12	1917	83.7	1304	12 ABL58447	ABL58447 Human pel
13	1917	83.7	3222	6 ABL58447	ABL58447 Human pel
14	1917	83.7	3526	6 ABL58447	ABL58447 Human pel
15	1917	83.7	7136	12 ADP48664	ADP48664 Human pel
16	1917	83.7	7136	13 ADP48664	ADP48664 Human pel
17	1917	83.7	7136	13 ADP48664	ADP48664 Human pel
18	1917	83.7	7136	14 ADP48664	ADP48664 Human pel
19	1917	83.7	7136	14 ADP48664	ADP48664 Human pel
20	1917	83.7	7136	14 ADP48664	ADP48664 Human pel
21	1912	83.5	1257	6 ABL58446	ABL58446 Murine pe
22	1912	83.5	1257	12 ADP48664	ADP48664 Mouse pel
23	1822	79.6	3846	6 ABL58452	ABL58452 Human pel
24	1642	71.7	1338	12 ADP48674	ADP48674 Human pel
25	1642	71.7	1338	12 ADP48674	ADP48674 Human pel
26	1642	71.7	2683	10 ADP48674	ADP48674 Human pel
27	1640	71.6	2589	6 ABL58452	ABL58452 Human pel
28	1633	71.3	2588	10 ADC30836	ADC30836 Human nov
29	1632	71.3	1936	6 ABL58452	ABL58452 Human pel
30	1618	70.7	3254	6 ABL58452	ABL58452 Human pel
31	1605	70.1	3332	3 ABL58452	ABL58452 Human sec
32	1574	68.7	2866	4 ABL58452	ABL58452 Human sec
33	1574	68.7	2866	4 ABL58452	ABL58452 Human sec
34	1482.5	64.7	2508	10 ABL58452	ABL58452 Human sec
35	1370	59.8	2183	4 ABL58452	ABL58452 Human sec
36	1370	59.8	2183	12 ABL58452	ABL58452 Human sec
37	1302	56.9	2966	4 ABL58452	ABL58452 Human sec
38	1192	51.2	16092	4 ABL58452	ABL58452 Human sec
39	1172	51.2	803	4 ABL58452	ABL58452 Human sec
40	1102	48.1	592	10 ADC32115	ADC32115 Human nov
41	1075	46.9	224	5 ABL58452	ABL58452 Human sec
42	1064	46.5	567	6 ABL58452	ABL58452 Human sec
43	1050	45.9	570	6 ABL58452	ABL58452 Human sec
44	1022	44.6	541	12 ABL58452	ABL58452 Human sec
45	996	43.5	31882	4 ABL58452	ABL58452 Human sec

ALIGNMENTS

RESULT 1
ID AAL46323 standard; cDNA; 1565 BP.
AC AAL46323;
DT 19-JUL-2002 (first entry)
XX
DE Human M33 coding sequence SEQ ID NO: 15.
KW Neurodegenerative disease; M30; M31; M32; M33; stroke;
KW fragile X syndrome; Huntington's disease; Parkinson's disease;
KW Alzheimer's disease; multiple sclerosis; ovarian cancer;
KW neurodegeneration; immune disorder; autoimmune disease; allergy;
KW infection; leukaemia; inflammation; neuroprotection; cerebroprotective;
KW immunosuppressive; cytostatic; neotropic; antiparkinsonian; antiallergic;
KW virulence; antiinflammatory; gene; ss.
OS Homo sapiens.
PN NO200221138-A2.
PD 14-MAR-2002.
PF 07-SEP-2001; 2001MO-BP010366.
PR 07-SEP-2000; 2000US-00657479.

XX 10-APR-2003.
 PD 24-SEP-2002; 2002MO-US030474.
 XX 24-SEP-2001; 2001US-0324631P.
 XX (HYSE-) HYSEQ INC.
 PA Zhang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX WPI: 2003-371981/35.
 DR P-PSDB; ADC31180.
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 PS Claim 1; SEQ ID NO 291; 1185bp; English.
 XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition, kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC coding sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the invention
 CC (ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1823 BP; 425 A; 483 C; 530 G; 385 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 9.69e-200 Length: 1823
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-041-030-4 (1-420) x ADC30209 (1-1823)
 QY 1 MetPheSerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
 Db 265 ATGTTTCCCTGGCGAGGAAACACTGGCCCCCAATAAGAGCCAGTGAAATACGGG 324
 QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgLys 40

Db 325 GAGCTGTGGTGTCTCGGATACAAATGCTTACCCAAATGAGATAGAGGACGAGGAAA 384
 QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
 Db 385 AGTAGATTGGCCCTCTACAAAGCCGCCAAAGGCAAAATGATGTCAAAACCCAGCACCGTCCAT 444
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
 Db 445 GTGATATCCAGGCCCAAGGATCCAAAGCTATCAGCTGCAGAAAGGTCAACACATATATCC 504
 QY 81 TyrThrLeuSerArgAsnGlnThrValAlaValGlyTyrThrHisAspLysAspThrAsp 100
 Db 505 TACACTTTGGCAAGATCAGACTGTGTGTGTGATACACACATATAGATAGATAGAT 564
 QY 101 MetPheGlnValGlyArgSerThrLysSerProLysAspPheValThrAspThrIle 120
 Db 565 ATGTTTCAGGTGGGAGATCAACAGAAAGCCATATGACTTCGTGTTCACAGACAGATT 624
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 625 TCTGGAGCCGAGAACACGAGCGAAGCCAGATTCACAGAGCACATATCCAGGTTCCG 684
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIaGlyPhe 160
 Db 685 TGCAGATCTGTGTCCACAGAGATGAACCTTACACAGACGATATTCGCCCGGATTT 744
 QY 161 AspSerSerLysAsnIlePheLeuGlyGlyLysAlaAlaLysTrpLysAsnProAspGly 180
 Db 745 GACTCTTCCAAAACATATTTCTTGAGAAAGGACGAAAGTGAAAGAACCCGACGGC 804
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 Db 805 CACATGATGTGGCTCACTACTATATGCGCTGTGTATGATCATCAGAGGGGCTTCACC 864
 QY 201 GluGluSerGlnProGlyValTPArgGluIleSerValCysGlyAspValTyrThrLeu 220
 Db 865 GAGAGATCCAGGCCCGGGGTCTGGCGGAGATTTCTGTCTGTGAGATGTGTACACTTG 924
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
 Db 925 CGAAGAACACAGGTGGGCCGAGCAACAGAGAAAGCTGTGGAAGTGAGACCAAGTCTTG 984
 QY 241 GluAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
 Db 985 CAGAGCGGCTCCCTCATTTGACTGTGTGGGGCCACTCTCTCTGAGAACAGCAGATGGG 1044
 QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnIleAsnAlaIa 280
 Db 1045 CTTTTCATATCTCAACTCAAGACATAGAAAGCCCTCCGAGAGATTAAGCCGCC 1104
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
 Db 1105 CGGCTCAAGTCTCTGTGGGGCTCAACACCTTGCTTCCCGACATCAACAGAAAGAG 1164
 QY 301 ValValGluGluLysGlnProTPAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
 Db 1165 GTGTGAGAGGAGAAAGAGCCCTGGGATATCTCAGTGTGTGCCACATTCAGGTTCCAC 1224
 QY 321 AsnTrpGlyHisArgSerAspThrGlnAlaAsnGluArgLysGlyProMetCysArgThr 340
 Db 1225 AACCTGGGGCCATGTGAGTGCACGAGGCGCAAGAGGAGAGTGTCCATGTCCAGACT 1284
 QY 341 ValGlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
 Db 1285 GTGGGCCCTATGTGCTCTGTGGCTGTGAGAGGAGAGATTTATAGAGCGAGGA 1344
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlyLysSerAlaLysTyr 380
 Db 1345 CCGCCAACTATATCTTCACTCCCTGTGACACGTGTGTCTGGAGAAAGTCTGCAAAATAC 1404
 QY 381 TrpSerGlnIleProLeuProHisGlyTyrThrHisAlaPheHisAlaIaCysProPheCys 400

QY 361 ProProThriSaIaphethrProCySGlyHieValCySerSergIuYsSerAlaLyTyx 380
 DB 1581 CCGCCAACTCATCTTCACTCCCTGTGACAGTGTCTGGAGAGTCTGCAAAATAC 1640
 QY 381 TrpSerGlnIleProLeuProHISglYThriSaIaphethrSaIaIaCyProPheCyS 400
 DB 1641 TGGTCTCAGATCCCGTTGCTCTCATGGAATCATGCAATTCAGCGTGTCTTCTCTGT 1700
 QY 401 AlAThGlnLeuValGlyGlnAerCySileYsLeuIlePheGlnIlyProIleAsp 420
 DB 1701 GCTACACAGCTGGTGGGAGCAAACTGCATCAAAATTAATTTCCAGGTCCCAATTGAC 1760
 RESULT 4
 ACC42349
 ID ACC42349 standard; cDNA; 5579 BP.
 XX
 AC ACC42349;
 XX
 DT 22-MAY-2003 (first entry)
 XX
 DE Human MAP kinase cascade activator #59 cDNA.
 XX
 KW Human; Elki phosphorylation; Elki phosphorylation kinase; virucide;
 KW antiinflammatory; immunomodulator; cyostatic; antiallergic; anti-HIV;
 KW antineumatic; antiarthritic; antidiabetic; antisthmatic; gene therapy;
 KW inflammation; autoimmune disease; viral disease; cancer; diabetes;
 KW rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
 KW IGA nephritis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003008589-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 15-JUL-2002; 2002MO-JP007174.
 XX
 PR 18-JUL-2001; 2001JP-00218204.
 PR 31-AUG-2001; 2001JP-00263450.
 PR 21-JAN-2002; 2002JP-00012176.
 XX
 PA (ASAH) ASAH KASEI KOGYO KK.
 XX
 PI Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
 XX
 DR WPI; 2003-229582/22.
 XX
 PT P-PSDB; ABR41083.
 PT Elki phosphorylation-associated gene and its encoded protein with MAP
 PT kinase cascade effect, applicable in diagnosis of and developing drugs
 PT for e.g. inflammations, autoimmune diseases, viral diseases and cancer.
 XX
 PS Claim 4; Page 631-640; 762pp; Japanese.
 XX
 CC The invention relates to a novel purified protein having Elki
 CC phosphorylation activity and/or an activity of activating Elki
 CC phosphorylation kinase. A protein of the invention has antiinflammatory,
 CC immunomodulator, virucide, cyostatic, antiallergic, antineumatic,
 CC antirheumatic, antidiabetic, antisthmatic, and anti-HIV activity. The
 CC polynucleotides may have a use in gene therapy. The gene and its encoded
 CC protein are applicable in diagnosis of and developing drugs for e.g.
 CC inflammations, autoimmune diseases, viral diseases and cancer such as
 CC rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral
 CC hepatitis and IGA nephritis. The present sequence is used in the
 CC exemplification of the invention
 XX
 SQ Sequence 5579 BP; 1535 A; 1111 C; 1214 G; 1719 T; 0 U; 0 Other;

Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-041-030-4 (1-420) x ACC42349 (1-5579)
 QY 1 MetPheSerProGlyngIugIuHISCySaIaProAsnLySgIuProValLyTyx 20
 DB 177 ATGTTTCCCTCGGCGAGAGAACCTGCGCCCAATTAAGAGCAGTGAATACGGG 236
 QY 21 GluLeuValIleuGlyTyraGngIaIaLeuProAsnGlyAspArgGlyArgGly 40
 DB 237 GAGCTGTGGTGTGGGTATCAATGTGTCTTAACCAATGAGATAGAGAGCGAGAA 296
 QY 41 SerArgPheAlaLeuTyrlYsaGp-rolYsaIaAsnGlyValIyProSerThrValHis 60
 DB 297 AGTAGATTGGCTCTTACAGCGCCCAAGGCAAAATGGTGTCAAAACCGACCGTCAT 356
 QY 61 ValIleSerThrProGlnIaSerLySaIaIleSerCySileYsLeuIlePheSer 80
 DB 357 GTGATATCCAGCCCGAGGATCCAGGCTATCAGCTCGCAAGGTCAACAGATATATCC 416
 QY 81 TyrThrLeuSerArgaGngInThrValIaIaGlyTyThriSaPlySaPThrAsp 100
 DB 417 TACACTTGTCAAGGATCAGACTGTGTGTGAGTATACACATGATTAAGATACGAT 476
 QY 101 MetPheGlnValGlyArgSerThrgIuSerProIleAspPheValIalThraPThrIle 120
 DB 477 ATGTTTCAGGTGGGAGATCAACAGAAACCCATGACTTCCTGTGTCAACAGCAT 536
 QY 121 SerGlySerGlnAerThraPngIuIaGngIleThrgInSerThrIleSerArgPheAla 140
 DB 537 TCTGGAGCGCAAGAACGAGACAGGCGAGATCAACAGACCAATTCAGATTCGCC 596
 QY 141 CySaArgIleValCySaPArgaGngIuProTyThrAlaArgIlePheAlaIaGlyPhe 160
 DB 597 TGCAGATGCTGTGGCAGACGAAATGAACCTTACAGACGAGTATTCGCCCGGATTT 656
 QY 161 AspSerSerLySaenIlePheLeuGlyGlyValAlaIaLySTpLySaP-rolAspGly 180
 DB 657 GACTCTTCCAAAMAACTATTTCTTGAGAAAGGCGCAAGGAGGAAAAACCCCGACGCC 716
 QY 181 HisMetAspGlyLeuThrThraPngIuValIleValMetHisProlArgGlyGlyPheThr 200
 DB 717 CACATGAAATGGGTCACTACTAATGCGCTGTGTATGATATTCAGAGGGGCTTCACC 776
 QY 201 GluGlnSerGlnProGlyValITraPArgIuIleSerValCySgIyAspValTyThrLeu 220
 DB 777 GAGGATTCACAGCCCGGGGTCTGGCCGAGATCTGTCTGTGAGATGTGTACACTTG 836
 QY 221 ArgGlnThraArgSerAlaGngInaArgGlyLyLeuValGlnSerGlnThraenValLeu 240
 DB 837 CGAGAAACCAAGTCGCGCCAGCAACAGAAAGCTGTGAAAGTGAAGACCAACGTCCTG 896
 QY 241 GlnAspGlySerLeuIleAspLeuCySgIyAlaThrLeuLeuThraPrgThrAlaAspGly 260
 DB 897 CAGAGAGGCTCCCTCTTAATACCTGTGTGGGCGACCTCTCTCGAGAAACAGAGATGGG 956
 QY 261 LeuPheHisThrProThrgInLySileGlnAlaLeuArgGlnIuIleAsnAlaIa 280
 DB 957 CTTTTCATATCTCAACTCAGAAACACATAGAAAGCCCTCCGAGAGATTAACGCCGCC 1016
 QY 281 ArgProGlnCySProValGlyLeuAerThrLeuAlaPheProSerIleAsnArgLySgIu 300
 DB 1017 CGGCTCAGGTGCTGTGGGGCTCAACACCTGCGCTTCCCGCATCAACAGAAAGAG 1076
 QY 301 ValValGlnGlnLySgInProThraIaTyThrLeuSerCySgIyHisIleGlyTyThris 320
 DB 1077 GTGGTAGAGAGAGAGAGAGCCCTGGGCAATCTCAGTTGTGGCCACGTGACCGGTACAC 1136
 QY 321 AsnTrgIyHisArgSerAspThrgIuAlaAsnGlnArgGlyCySProMetCySaTyThr 340
 DB 1137 AACCTGGGCGCATGAGATACACAGAGCGCAACAGAGAGAGATGTCTCATGTGCAAGACT 1196

Alignment Scores:

Pred. No.: 4,46e-199
 Score: 2290.00
 Percent Similarity: 100.0%
 Best Local Similarity: 100.0%

Length:

5579
 Matches: 420
 Conservative: 0
 Mismatches: 0

QY 341 ValGlyProTyrValProLeuTyrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
Db 1197 GTGGGCCCCATGAGCTCTCTGCTTGCTGAGGAGAGATTATTATGAGAGCAGGA 1256
QY 361 ProProThrtHtHsAlaPheThrProCysGlyVhtValCysSerGluLysSerAlaLysTyr 380
Db 1257 CCGGCACACTGATGCTTTCACCTCCCTGGAGACCGTGGCTCGGAGAACTCGCAAAATAC 1316
QY 381 TrpSerGlnIlePheProLeuProHtIleGlyThrHtHsAlaPheHtHsAlaCysProPheCys 400
Db 1317 TGGTCTCAGATCCCGTTCCTCATGAGAACATCATGATTCACGCTGCTTGCCTTCTGT 1376
QY 401 AlaThrGlnLeuValGlyGlnGlnAsnGlyIleLysLeuIlePheGlnGlyProIleAsp 420
Db 1377 GCTACACAGCTGTTGGGAGCAAACTGCATCAATTAATTTCCAAAGTCCAAATTGAC 1436

RESULT 5
ABLS8449
ID ABL58449 standard; DNA; 1263 BP.
XX
AC ABL58449;
XX
XX 30-JUL-2002 (first entry)
XX
XX Human pellino-2 polypeptide coding sequence.
XX
XX Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide;
XX anti-bacterial; fungicide; protozoicide; antischistosomal; antirheumatic;
XX antiarthritic; antiinflammatory; antithrombotic; neuroprotective;
XX monocytic; anticancer; human; pellino-2; gene; ds.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1263
FT /*tag= a
FT /product= "pellino-2"
XX
XX WO200183739-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US013676.
XX
XX 28-APR-2000; 2000US-0200198P.
XX
XX (IMMV) IMMUNEX CORP.
XX
XX Bird TA, Cosman DJ;
XX
XX WPI; 2002-066532/09.
XX
XX P-PSDB; ABB07922.
XX
XX
XX New pellino polypeptides for identifying compounds that alter polypeptide
XX activity, treating pathogenic infection or inhibiting apoptosis, are
XX capable of stimulating nuclear factor-kappaB- or p38-dependent
XX transcription.
XX
XX
XX Claim 3; Page 61; 70pp; English.
XX
XX
XX The invention provides polypeptides capable of stimulating nuclear factor
XX (NF)-kappaB-dependent transcription or p38-dependent transcription,
XX referred as pellino polypeptides. The pellino polypeptides are useful for
XX identifying modulators that alter the pellino polypeptide and pellino
XX dominant-negative activity. They are also useful for identifying
XX compounds that inhibit the binding activity of the polypeptides and to
XX study cell-signal transduction. They are useful for preventing or
XX treating infection by a pathogen such as virus, bacterial, fungi, algae
XX or protozoa, or inhibiting apoptosis. Dominant-negative pellino
XX polypeptides are useful for treating inflammatory conditions such as
XX asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's
XX disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and
XX also for inhibiting mitogen activated protein (MAP) kinase-activated

CC pathways. Pellino polypeptides and polymucleotides are useful to identify
CC small molecule inhibitors of protein association or function of pellino,
CC and other molecules involved in interleukin (IL)-1 signaling. The present
CC sequence represents a human pellino-2 polypeptide coding sequence
XX
SQ Sequence 1263 BP; 317 A; 336 C; 351 G; 259 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,576-199 Length: 1263
Score: 2283.00 Matches: 419
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: Gaps: 0
US-10-041-030-4 (1-420) x ABL58449 (1-1263)
QY 1 MetPheSerProGlyGlnGlnGlnHtHsCysAlaProAsnLysGluProValLysTyrGly 20
Db 1 ATGTTTCCCTGGCCAGAGAACTCGGCCCAATMGAGAGCCAGTAATACGGG 60
QY 21 GluLeuValValLeuGlyTyrAsnGlyValAlaLeuProAsnGlyAspArgGlyArgArgLys 40
Db 61 GAGCTGGTGGTCTCGGATCAATGGTGTTCACCAATGAGATGAGAGCGAGGAAA 120
QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHs 60
Db 121 AGTAGATTCCTCCCTTCAAGCGGCCCAAGGCAAAATGTTCAACCCAGCACCGTCCAT 180
QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysValGlyGlnHtHsSer 80
Db 181 GTGATATCCAGCCCGAGGATCCAAAGGCTATCAGCTGCAAAAGTCAACACATATATCC 240
QY 81 TyrThrLeuSerArgAsnGlnThrValValValGlnTyrThrHtHsPheLysAspThrAsp 100
Db 241 TACACTTTGCAAGGATCACTGCTGTGGTGGATGACACATGATTAAGATACGGAT 300
QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValValThrAspThrIle 120
Db 301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGACTTCTTGTCAACAGACAGATT 360
QY 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCTGGCAGCCAGAAACGAGCAGAACCCAGATCACACAGACACATATCCAGTTGCGC 420
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
Db 421 TGCAGATTCGTGCGACAGAAATGAACCTTACACAGCAGCATATTCGCGCGGATTT 480
QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGly 180
Db 481 GACTCTTCCAAAACATATTTCTTGGAGTAAAGCAGCAAAAGTGAAAAAAGCCGACGCG 540
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 541 CACATGATGGGCTCACTACTAATGCGTCTGTGATGATGCACAGAGGGGCTTCACC 600
QY 201 GluGlnSerGlnProGlyValLTrpArgGlnLysSerValCysGlyLysPheValTyrThrLeu 220
Db 601 GAGAGATCCAGCCCGGGGTCTGGCCGAGATCTGTCTGTGAGATGTCACCTTG 660
QY 221 ArgGlnThrArgSerAlaGlnGlnArgGlyLysLeuValGlnSerGluThrAsnValLeu 240
Db 661 CGAGAAACAGGTCTGGCCCGAGAACAGAGAAAGCTGTGAAAGTGAAGCAACGCTCTG 720
QY 241 GlnAspGlySerLeuIleAspLeuGlyValAlaThrLeuLeuTrpArgThrAlaAspGly 260
Db 721 CAGGAGGGTCCCTCATTTAGCTGTGTGGGGCACTCTCTCTGGAAGACAGAGATGGG 780
QY 261 LeuPheHisThrProThrGlnLysHisIleGlnAlaLeuArgGlnGlnLysAlaAla 280
Db 781 CTTTTCATCTCACTCAACTCAGAAAGCATGAAAGCCCTCGGACAGGAGATTAACGCGGCC 840

QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
 DB 841 CGGCTCAGTCTCTCTGGGGGCTCAACCTGGCTTCCCGACATCAACAGGAAGAG 900
 QY 301 ValValGlnGluValGlnProTTPalaTyrLeuSerCysGlyValHisGlyTyrHis 320
 DB 901 GTGGTGGAGAGAGACAGCCCTGGGCAATCTCATGTTGTGGCCACGTGACAGGGTACAC 960
 QY 321 AsnTTPGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
 DB 961 AACTGGGGCCATCGAGATGACACGGAGCCACAGAGGGAGTGTCCCATGTGACACT 1020
 QY 341 ValGlyProTyrValProLeuTTPLeuGlyCysGlyValGlyPheTyrValAspAlaGly 360
 DB 1021 GTGGGGCCCTATGTGCTCTCTCTGGCTTGGCTGAGGAGAGATTATTTATGACCGAGGA 1080
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlyValSerAlaTyr 380
 DB 1081 CCGGCAACTCATGCTTTCACCTCCCTGGACAGGTGTCCGAGAGAGTCCGAGAAATAC 1140
 QY 381 TTPSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
 DB 1141 TGGTCTCAGATCCCGTGGCTCATGGAATCATGATGATTCACGCTGCTTCCCTTCTGT 1200
 QY 401 AlaThrGlnLeuValGlyGlnGluAsnCysAlleValLeuIlePheGlnGlyProIleAsp 420
 DB 1201 GCTACACAGCTGTGGTGGAGCAAACTGCATCAATTAATTTCCAAAGTCCCAATGAC 1260
 RESULT 6
 ADP48670 ID ADP48670 standard; cDNA; 1263 BP.
 AC ADP48670;
 XX 09-SEP-2004 (first entry)
 DT Human Pellino-2 encoding cDNA SEQ ID NO:7.
 XX
 XX nuclear factor kappa B dependent transcription inhibitor;
 KW NF-kB-dependent transcription inhibitor;
 KW p38-dependent transcription inhibitor; Pellino-1;
 KW interleukin 1 receptor-associated kinase 4;
 KW IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
 KW antiarthritic; antiasthmatic; antiinflammatory; antirheumatic; antiulcer;
 KW gastrointestinal; neuroprotective; nootropic; IL-1 antagonist;
 KW IRAK protein kinase family inhibitor;
 KW NF kappa B inducing kinase inhibitor; asthma; rheumatoid arthritis;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW atherosclerosis; Alzheimer's disease; human; Pellino-2; chromosome 14;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1263
 FT /*tag= a
 FT /product= "Pellino-2"
 XX
 XX MO2004053092-A2.
 XX
 XX 24-JUN-2004.
 XX
 XX 09-DEC-2003; 2003WO-US039188.
 XX
 XX 11-DEC-2002; 2002US-00317250.
 XX
 XX (IMMV) IMMOTEX CORP.
 XX (CLEV-) CLEVELAND CLINIC FOUND.
 XX Bird TA, Cosman DJ, Li X;
 XX WPI: 2004-480927/45.
 XX P-PDB: ADP48671.

XX
 PT Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
 PT asthma, and rheumatoid arthritis, comprises assaying the association of a
 PT Pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
 PS Claim 9; SEQ ID NO 7; 81bp; English.
 XX
 CC The present invention describes a method for identifying compounds (C)
 CC that inhibit nuclear factor kappa B (NF-kB)-dependent transcription or
 CC p38-dependent transcription. The method comprises: (1) mixing a test
 CC compound with a Pellino-1 polypeptide; (b) assaying the association of
 CC the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4)
 CC in the presence of the test compound; and (c) determining whether the
 CC test compound inhibits the association of the Pellino-1 polypeptide with
 CC a binding partner. Also described: (1) an inhibitory nucleic acid that
 CC binds to a nucleic acid encoding an amino acid sequence as described
 CC above, where the presence of the inhibitory nucleic acid within a cell
 CC inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory
 CC polypeptide comprising an antibody fragment that binds to a polypeptide
 CC comprising an amino acid sequence as described above, where the presence
 CC of the inhibitory polypeptide within a cell inhibits the association of
 CC Pellino-1 with IRAK-4. (C) have antiarteriosclerotic, antiarthritic,
 CC antiasthmatic, antiinflammatory, antirheumatic, antiulcer,
 CC gastrointestinal, neuroprotective and nootropic activities, and can be
 CC used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF
 CC kappa B inducing kinase inhibitor. The methods and compositions of the
 CC present invention are useful for the prevention and/or treatment of
 CC diseases or conditions associated with aberrant expression or activity of
 CC the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC atherosclerosis and Alzheimer's disease. The present sequence encodes
 CC human Pellino-2, which is used in the exemplification of the present
 CC invention. The human Pellino-2 gene is located on chromosome 14, more
 CC specifically to 14q24.3.
 XX
 SQ Sequence 1263 BP; 317 A; 336 C; 351 G; 259 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,57e-199 Length: 1263
 Score: 2283.00 Matches: 419
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 99.7% Indels: 0
 DB: 12 Gaps: 0
 US-10-041-030-4 (1-420) x ADP48670 (1-1263)
 QY 1 MetPheSerProGlnGlnGluGlnHisCysAlaProAsnIleGluProValIleTyrGly 20
 DB 1 ATGTTTTCCTCCCTGGCCAGAGGAGAACCTGGCCCAATAGAGAGCCAGTGAATACGGG 60
 QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyValAspArgIleArgTglys 40
 DB 61 GAGCTGTGTGTCTGGGTATCATGTGCTTTACCAATGAGATAGAGAGCGAGAGAAA 120
 QY 41 SerArgPheAlaLeuTyrIleValArgProIleValAsnGlyValIleAspSerThrValHis 60
 DB 121 AGTAGATTGGCCCTCTCAAGCGCGCCCAAGGCAAAATGTGTCAAAACCGACACGCTTCAT 180
 QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCysValGlyGlnHisSerIleSer 80
 DB 181 GTGATATCCAGCCCAAGGATCCAAAGGCTATAGCTGCAAAAGTCAACACAGATATATCC 240
 QY 81 TyrThrLeuSerArgAsnGlnThrValValValGlyTyrThrHisAspIleAspThrAsp 100
 DB 241 TACACTTTGTCAAGATCAAGATGAGTGTGTGTGATGATCAACATATAGATAGAGAT 300
 QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
 DB 301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGATGATCTGTGTCAACAGAGATT 360
 QY 121 SerGlySerGlnAsnThrAspGluValGlnIleThrGlnSerThrIleSerArgPheAla 140

Dp	361	TCCTGGACGACCAGAACACGGAGCGAAGCCCAATACACAGAGACACCATATCCAGGTTGGCC	420
Oy	141	CyAaTgllEvaLcYabPaTgAbmGluProTyThrAlaArgllEPhelAlaGlyPhe	160
Dp	421	TGCAGGAATCGTGTGGACAGAAATGAACTTACACAGACGGAATATTCGCGCGGATTT	480
Oy	161	AaPseSerLysAaenLlePheLeuGlyGlyLulysAlaAlaLysTPlyAaPProAaPgly	180
Dp	481	GACCTTCCAAAAACATATTTCTTGAGAGTAAGGACGAGAAAGTGAANAACCCGACGGC	540
Oy	181	HisMeaAaPglyLeuThrThrAaenGlyValLeuValMeChsProAaPglyLysPheThr	200
Dp	541	CACATGATGGGCTCACTACTAATGCGCTCCGTGGATGATCATCCACGAGGGGGCTTACCC	600
Oy	201	GlulGusSerGlnProGlyValTPPaArgLulLeSerValCybGlyAaPValTYrThrLeu	220
Dp	601	GAGAGATCCACAGCCCGGGGTCTGGGCGAAGATCTCTGTGTGAGATGTTACACCTTG	660
Oy	221	ArgGluThrAaGSeLaGlnGlnArgGlyLysLeuValGluSeGluThrAaValLeu	240
Dp	661	CGAGAAACCAAGTGGGCCACCAACAGAGAAAGCTGGTGGAAAGTGAACCAAGTCTCTG	720
Oy	241	GlnAaPglySerLeuLleAaPLeuCybGlyAlaThrLeuLeuTPPaGThrAlaAaPgly	260
Dp	721	CAGACGCGCTCCCTCATATGACCTGTGTGGGGCCACTCTCTCTGGAGAACAGCAATGGG	780
Oy	261	LeuPheHsThrProThrGlnLysHsAlleGlnLalaLeuAaGnglnLlAaAaAla	280
Dp	781	CTTTTTCATACTCCAACTCGAAGACATGAAGAGCCCTCCGCGAGAAATTTAAGCGGCC	840
Oy	281	ArgProGlnCybProValGlyLeuAaThrLeuAlaPheProSerLleAaAaGlyAaGlu	300
Dp	841	CGGCTCAGTGTCTCTGTGGGGCTCAACACCTGGGCTTCCCAACATCAACAGAAAGAG	900
Oy	301	ValValGlnGlnLysGlnProThrAlaTyrlLeuSerCybGlyHsValHsGlyTYrHs	320
Dp	901	GTTGTGGAGAAAGACAGCCCTGGGCATATCTCAAGTTGTGGCACGTGACAGGGATACAC	960
Oy	321	AaenTPrgLysHsArgSerAaPThrGlnLalaAaGlnAaArgGluCybProMetCybArgThr	340
Dp	961	AACAGGGGCCATCCGAGTGAACACGAGGCCAACAGAGAGAGTGTCCCATGTGCAGGACT	1020
Oy	341	ValGlyProTyValProLeuTPleuGlyCybGlyAlaGlyPheTyValAaPAlaGly	360
Dp	1021	GTTGGGCCCCATATGGCTCTCTGTGGTGTGGCTGTGAGGACGAGATTTTATGTGAGCGCAGA	1080
Oy	361	ProProThrHsAlaPheThrProCybGlyHsValCybSerGlnLysSerAlaTyPTr	380
Dp	1081	CCGCCAACTATGCTTTCACTCCCTGTGGACACGTGTCTCGAGAAAGTCTGCAAAATAC	1140
Oy	381	TPSerGlnLleProLeuProHsGlyThrHsAlaPheHsAlaAlaCybProPheCyb	400
Dp	1141	TGATCTCAGATCCCGTTGCTCTCATGAACTCATGATTCACGCTGCTTGCCTTTGCT	1200
Oy	401	AlaThrGlnLeuValGlyGlnGlnAaenCybLysLeuLlePheGlnGlyProLleAaP	420
Dp	1201	GCTACACGCTGTGTGGGAGCAAAATGCATCAATTAATTTTCCAGAGTCCAAATGAC	1260

RESULT 7
 ABL58448
 ID ABL58448 standard; DNA, 1260 BP.
 XX
 XX ABL58448;
 AC
 XX
 DT 30-JUL-2002 (first entry)
 XX
 XX Murine pellino-2 coding sequence.
 DE
 XX
 XX
 XX Nucleic factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide;
 KM antibacterial; fungicide; protozoacide; antiasthmatic; antirheumatic;
 KM antiatheritic; antiinflammatory; antithrombotic; neuroprotective;
 KM neurotropic; antitumor; mouse; pellino-2; gene; ds.
 XX
 XX

```

OS      Mus musculus.
XX      Key      Location/Qualifiers
XX      FH      1. .1257
XX      CDS
XX      FT      /*tag= a
XX      FT      /product= "pellino-2"
XX      MN      M0200183739-A2.
XX      PD      08-NOV-2001.
XX      PE      27-APR-2001; 2001MO-US013676.
XX      PR      28-APR-2000; 2000US-0200198P.
XX      PA      (IMMIV ) IMMUNEX CORP.
XX      PI      Bird TA, Cosman DJ;
XX      DR      WPI; 2002-066532/09.
XX      DR      P-PSDB; ABB07921.
XX      PT      New Pellino polypeptides for identifying compounds that alter polypeptide
XX      PT      activity, treating pathogenic infection or inhibiting apoptosis, are
XX      PT      capable of stimulating nuclear factor-kappaB- or p38-dependent
XX      PT      transcription.
XX      PS      Example; Page 58-59; 70pp; English.
XX      CC      The invention provides polypeptides capable of stimulating nuclear factor
XX      CC      (NF)-kappaB-dependent transcription or p38-dependent transcription,
XX      CC      referred as pellino polypeptides. The pellino polypeptides are useful for
XX      CC      identifying modulators that alter the pellino polypeptide and pellino
XX      CC      dominant-negative activity. They are also useful for identifying
XX      CC      compounds that inhibit the binding activity of the polypeptides and to
XX      CC      study cell-signal transduction. They are useful for preventing or
XX      CC      treating infection by a pathogen such as virus, bacterial, fungi, algae
XX      CC      or protozoa, or inhibiting apoptosis. Dominant-negative pellino
XX      CC      polypeptides are useful for treating inflammatory conditions such as
XX      CC      asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's
XX      CC      disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and
XX      CC      also for inhibiting mitogen activated protein (MAP) kinase-activated
XX      CC      pathways. Pellino polypeptides and polynucleotides are useful to identify
XX      CC      small molecule inhibitors of protein association or function of Pellino,
XX      CC      and other molecules involved in interleukin (IL)-1 signaling. The present
XX      CC      sequence represents a murine pellino-2 polypeptide coding sequence
XX      SQ      Sequence 1260 BP; 309 A; 357 C; 359 G; 235 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1.03e-189      Length:      1260
Score:      2175.50      Matches:      399
Percent Similarity:      97.4%      Mismatches:      10
Best Local Similarity:      95.0%      Indels:      1
Query Match:      95.0%      Gaps:      1

US-10-041-030-4 (1-420) x ABL58448 (1-1260)

QY      1 MetPheSerProGlyGlnGlnGlnLHnISCySaLaPProAenLyvGIuPProValLyvTYTGly 20
Db      1 ATGTGTTTCCCGGCGCAGAGAGAACCCAGCGCGCCCAACAGGAGCGCGTGAATACGGG 60
QY      21 GluLeuValValIleuGlyTYvRaNGlyValaLeuPProaNGlyvAaPaRvGGLvARvARgLyv 40
Db      61 GAGGTGGTGGTCTCTGGGGTACATAGTGTCTTACTTAATGGTGAACAGGGGCGAGAGAAA 120
QY      41 SerARvPheLaLeuTYvLySaRvProLyvAlaNGlyValLyvProSerThValHis 60
Db      121 AGCAGATTGGCCCTCTATAGCGGACCTACGCAAGTGGTGTCAAAACCAAGACAATCCAC 180
QY      61 ValIleSerThProGlnAlaSerLyvAlaIleSerCyvLyvGlyGlnHisSerIleSer 80

```


Db 181 ATGGTCTCCACACACAGCGCTCCAGCCATGAGCTCCAGAGGACATCAGCATATCG 240
 QY TyrThrLeuSerArgangInThrValValGluTyrThrHisAspIysAspThrAsp 100
 Db 241 TACACGTTGTCACGAGCGAGCGAGCGTAGTGATGATACACACGATTAAGACACCGAC 300
 QY MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
 Db 301 ATCTTTACAGTGGGAGAGTCAACAGAAAGCCCTTATCATCTTGGTGTCAACAGACCGGTT 360
 QY SerGlySerGlnanthraAspGluValGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 361 TCCGCGCGGTGACAGACGAAAGAT---GCCCAATGATCAACAGAGCAACATCTCTAGGTTGCA 417
 QY CysArgIleValCysAspArgangInuProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 Db 418 TGCAGAGTCGTGTGTACAGAAACGAGCCATATACAGACGCAATTCGCGGCGAGATTC 477
 QY AspSerSerIysAsnIlePheLeuGlyGluValAlaIleValTpyIAsnProAspGly 180
 Db 478 GATTTCTCCAAAATATCTTTCTTGAGAGAAAGACCAAAATGGAATAACCTTGATGGA 537
 QY HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgIleGlyPheThr 200
 Db 538 CATATGATGAGTCACTACCAATGATGTCCTAGTATGACCCGCAAGAGGAGCTTACCC 597
 QY GluGluSerGlnProGlyValITrpaArgIleIleSerValCysGlyAspValITyrThrLeu 220
 Db 598 GAGGAATCCAGCGCTGAGTGTGAGAGAGATCTCTGTGTGGGAGTGTGTACACCTTG 657
 QY ArgGluThrArgSerIleGlnIleArgGlyIleLeuValGluSerGluThrAsnValLeu 240
 Db 658 CGAGAGCCAGGTGCGGCCCGACAGAGGGGAAAGCTGTGGAAGGTGAGCAACAGTCTTG 717
 QY GlnAspGlySerLeuIleAspLeuCysGlyValAThrLeuLeuThrArgThrAlaAspGly 260
 Db 718 CAGACGGGCTCCCTCATTTGACCTGTGTGGGGCCACTCTCTCTGGAACACCGAGATGGC 777
 QY LeuPheHisThrProThrGlnIleValHisIleGluAlaLeuArgGlnIleValAlaAla 280
 Db 778 CTTTTCACCCCTCTACTCAGAAACATAGAAAGCCCTCCGCGAGAGATCAATGCGAGCC 837
 QY ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgValGlu 300
 Db 838 CGAGCCCAAGTGGCCCGGGGCTTTAACACCTGTGCTTCCCGACATCAACCGAAGGAA 897
 QY ValValGluGluValGluProTyrPalaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
 Db 898 GTGGTGGAAAGAAAGACGCCCTGGGCACTACCTGAGCTGGGCAATGTGCACGGCTTACAC 957
 QY AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgIleCysProMetCysArgThr 340
 Db 958 AGCTGGGGGCATCGAGCGAGCGAGGAGCAAGCAAGAGGAGGAGTGTCCCATGTGAGAGACT 1017
 QY ValGlyProTyrValProLeuTrpLeuGlyCysGluAlaIleGlyPheTyrValAspAlaGly 360
 Db 1018 GTGGGGCCCTTACCTCTCTGTGTGGGTGTGAGGACAGATTTTATGTGCAATGCGGGA 1077
 QY ProProThrHisAlaPheThrProCysGlyHisValValCysSerGluIysSerAlaValTyr 380
 Db 1078 CCCCACATCAACGCTTACCCCTTGGGAGAGTCTGTTCAGAAAAGTCTGCCAAGTAC 1137
 QY TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
 Db 1138 TGTGTCCAGATCCACATGCGCCACGAGAACGCAACGCTTATGCGCTGTGCG 1197
 QY AlaThrGlnLeuValGlyGluGlnAsnCysIleIleLeuIlePheGlnGlyProIleAsp 420
 Db 1198 GCCACGACGCTGTGTGTGAACGAACTGCATCAATTTGATTTTCCAAAGGTCCAGTGGAC 1257

RESULT 8
 ADP48668
 ID ADP48668 standard; cDNA; 1260 BP.

XX AC ADP48668;
 XX DT 09-SEP-2004 (first entry)
 XX DE Mouse Pellino-2 encoding cDNA SEQ ID NO:5.
 XX KW nuclear factor kappa B dependent transcription inhibitor;
 KW NF-kB-dependent transcription inhibitor; Pellino-1;
 KW p38-dependent transcription inhibitor; Pellino-1;
 KW interleukin 1 receptor-associated kinase 4;
 KW IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
 KW antiarthritic; antiasthmatic; antiinflammatory; antirheumatic; antitumor;
 KW gastrointestinal, neuroprotective; nootropic; IL-1 antagonist;
 KW IRAK protein kinase family inhibitor;
 KW NF kappa B inducing kinase inhibitor; asthma; rheumatoid arthritis;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW atherosclerosis; Alzheimer's disease; mouse; Pellino-2; gene; ss.
 XX OS Mus musculus.
 XX FH Key
 XX FT CDS 1..1260
 XX FT /*tag= a
 XX FT /product= "Pellino-2"
 XX MO2004053092-A2.
 XX PD 24-JUN-2004.
 XX PF 09-DEC-2003; 2003MO-US039188.
 XX PR 11-DEC-2002; 2002US-00317250.
 XX PA (IMMUNEX CORP.
 XX PA (CLEV-) CLEVELAND CLINIC FOUND.
 XX PI Bird TA, Cosman DJ, Li X;
 XX WPI; 2004-480927/45.
 XX DR P-PSDB; ADP48669.
 XX PT Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
 PT acthina, and rheumatoid arthritis, comprises assaying the association of a
 PT Pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
 XX PS Example 1; SEQ ID NO 5; 81pp; English.
 XX The present invention describes a method for identifying compounds (C)
 CC that inhibit nuclear factor kappa B (NF-kB)-dependent transcription or
 CC p38-dependent transcription. The method comprises: (1) mixing a test
 CC compound with a Pellino-1 polypeptide; (b) assaying the association of
 CC the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4)
 CC in the presence of the test compound; and (c) determining whether the
 CC test compound inhibits the association of the Pellino-1 polypeptide with
 CC a binding partner. Also described: (1) an inhibitory nucleic acid that
 CC binds to a nucleic acid encoding an amino acid sequence as described
 CC above, where the presence of the inhibitory nucleic acid within a cell
 CC inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory
 CC polypeptide comprising an antibody fragment that binds to a polypeptide
 CC comprising an amino acid sequence as described above, where the presence
 CC of the inhibitory polypeptide within a cell inhibits the association of
 CC Pellino-1 with IRAK-4. (C) have antiarteriosclerotic, antitumor,
 CC antiasthmatic, antiinflammatory, antirheumatic, antitumor,
 CC gastrointestinal, neuroprotective and nootropic activities, and can be
 CC used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF
 CC kappa B inducing kinase inhibitor. The methods and compositions of the
 CC present invention are useful for the prevention and/or treatment of
 CC diseases or conditions associated with aberrant expression or activity of
 CC the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC atherosclerosis and Alzheimer's disease. The present sequence encodes
 CC mouse Pellino-2, which is used in the exemplification of the present

CC invention.
 XX Sequence 1260 BP; 309 A; 357 C; 359 G; 235 T; 0 U; 0 Other;
 SQ Alignment Scores:
 Pred. No.: 1,83e-189 Length: 1260
 Score: 2175.50 Matches: 399
 Percent Similarity: 97.4% Conservative: 10
 Best Local Similarity: 95.0% Mismatches: 10
 Query Match: 95.0% Indels: 1
 DB: 12 Gaps: 1
 US-10-041-030-4 (1-420) x ADPA8668 (1-1260)
 QY 1 MetPheSerProGluGlnGluGlnHlaCysAlaProkenuylsGluProValLysTYrGly 20
 DB 1 ATTTTTCCTCCGGGCGGAGAGAACCCAGCGCCCCCAAGAGACCGGTGAAATACGGG 60
 QY 21 GluLeuValValLeuGlyTYrAenglyAlaLeuProkenuylsAparGlyAArgArgLys 40
 DB 61 GAGCTGTGGTCTCGGGGTACAAATGGTCTTACCTTAATGTGACAGGGGCAAGAGGAAA 120
 QY 41 SerArgPheAlaLeuTYrLysAArgProLysAlaAenglyValLysProSerThrValHis 60
 DB 121 AGCAGATTGGCCCTATTAAGCCGACCTACGCTGTCTCAAAACCCACACATTCAC 180
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
 DB 181 ATGGTCTCCACACACAGCGGCTCCAGGCCATGCTCCAGGACATCAACGCAATATCG 240
 QY 81 TYrThrLeuSerArgAenglnThrValValValGlyTYrThrHisAerLysAerThrAsp 100
 DB 241 TACACGTTGTCAAGAGCCAGACGAGTAGTGTGTGAAGACACACAGTAAGAAGACGGAC 300
 QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAerPheValValThrAerThrIle 120
 DB 301 ATTTTCAGGTGGGCGAGGTCAACAGAAAGCCCATTCATCTCGGTGCACAGACACCGTT 360
 QY 121 SerGlySerGlnAerThrAerGluValGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 361 TCCGGCGGTGAGAAAGAAAT--GCCCAATCAACAGAGCACCATCTTACGTTGCA 417
 QY 141 CysArgIleValCysAerAerAenglnProTYrThrAlaArgIlePheAlaIleGlyPhe 160
 DB 418 TGCAGAGATCGTGTGTGACAGAAAGACCATATACACAGCAATATCGGGCAGATTC 477
 QY 161 AepSerSerLysAerIlePheLeuGlyGluValAlaIleTYrLysAerProAerGly 180
 DB 478 GATTCTTCCAAAATATCTTCTTGGAGAAAGCACAAATGGAATGCCATGATGGA 537
 QY 181 HisMetAerGlyLeuThrThrAenglyValLeuValMetHisProAerGlyGlyPheThr 200
 DB 538 CACATGATGACACTACATCAATGAGTCTTGTGTGTGACCCGCAAGAGGCTTACC 597
 QY 201 GluGluSerGlnProGlyValITPArgGluIleSerValCysGlyAerValTYrThrLeu 220
 DB 598 GAGGAATCCAGAGCTGAGATCTGAGAGAGATCTCTGTGTGGAGTGTGTACACTTG 657
 QY 221 ArgGluThrArgSerAlaGlnGlnAlaGlyValLeuValGluSerGluThrAerValLeu 240
 DB 658 CGAGAGACCAAGTCCGCGCCAGAGAGGGGAAAGCTGTGGAAGTGAACAACAGTCTTG 717
 QY 241 GluAerGlySerLeuIleAerLeuCysGlyAlaThrLeuLeuThrAerGluIleAerGly 260
 DB 718 CAAGACGGCTCCCTCATTTGACCTGTGTGGGCCCTCTCTGTGGAACCGCAGATGGC 777
 QY 261 LeuPheIleThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAerAla 280
 DB 778 CTTTTCACGCTCTACTACAGAAACATAGAACCTCCGGCAGAGATCATATGCGACC 837
 QY 281 ArgProGlnCysProValGlyLeuAerThrLeuAlaPheProSerIleAerArgLysGlu 300
 DB 838 CGACCCAGATGGCCGTTGAGGCTTAAACACCTGACCTTCCCAAGCATCAACGGAGAGAA 897

QY 301 ValValGluGluLysGlnProThrAlaTYrLeuSerCysGlyValValHisGlyTYrHis 320
 DB 898 GTGTGGAGAGAGAGACGCTCTGGGCATACCTAGCTGGCCCATGTGCACGGCTACAC 957
 QY 321 AenTYrGlyHisAerSerAerThrGluAlaAenglyValGluCysProMetCysAerGlyThr 340
 DB 958 AGCTGGGGCCATGTGGAGCGCAGCGGAAGCCAAAGAGAGAGATGTCCTCATGTGCAGACT 1017
 QY 341 ValGlyProTYrValProLeuThrLeuGlyCysGluAlaGlyPheTYrValAerAlaGly 360
 DB 1018 GTGGGCCCTACCTCTCTGTGGCTGGGCTGTGAAGCAGAAATTTATGTGATGCGAGGA 1077
 QY 361 ProProThrHisAlaPheThrProCysGlyValValCysSerGluLysSerAlaLysTYr 380
 DB 1078 CCCCACACTACGCTTACACCCCTCGGGCAGCTGTGTTCAGAAAAGTGTGCAGATAC 1137
 QY 381 TrpSerGlnIleProLeuProHisGlyTYrThrHisAlaPheHisAlaLysProPheCys 400
 DB 1138 TGGTCCAGATCCCATCTGCCCCCAGGACGACAGCGCTTCATCCCGCTGTCTGC 1197
 QY 401 AlaThrGlnLeuValGlyGluGlnAerCysIleLeuLeuIlePheGlnGlyProIleAer 420
 DB 1198 GCCACGACGTGGTGTGTGAACGAATGCAATTAATTTCCAAAGGTCCAGTGAC 1257
 RESULT 9
 ID ACC42348 standard; cDNA; 1717 BP.
 XX ACC42348;
 AC 22-MAY-2003 (first entry)
 DT 22-MAY-2003 (first entry)
 XX
 DE Mouse MAP kinase cascade activator #9 cDNA.
 XX
 XX Mouse; Elki phosphorylation; Elki phosphorylation kinase; virucide;
 KW antiinflammatory; immunomodulator; cyostatic; anti-allergic; anti-HIV;
 KW antidiabetic; antidiabetic; antidiabetic; antidiabetic; cancer therapy;
 KW inflammation; autoimmune disease; viral disease; cancer; diabetes;
 KW rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
 KW Iga nephritis; gene; ss.
 XX
 XX Mus musculus.
 OS
 XX NC02003008589-A1.
 PN
 XX 30-JAN-2003.
 PD
 XX 15-JUL-2002; 2002MO-JP007174.
 PF
 XX 18-JUL-2001; 2001JP-00216204.
 PR 31-AUG-2001; 2001JP-00263450.
 PR 21-JAN-2002; 2002JP-00012176.
 XX
 PA (ASAH) ASAH KASEI KOGYO KK.
 PI Matsuzaki O, Matsuda A, Nagano Y, Suzuki N,
 XX WPI; 2003-229582/22.
 DR P-PSDB; ABR41082.
 DR
 XX
 PT Elki phosphorylation-associated gene and its encoded protein with MAP
 PT kinase cascade effect, applicable in diagnosis of and developing drugs
 PT for e.g. inflammation, autoimmune diseases, viral diseases and cancer.
 XX
 PS Claim 4; Page 623-627; 762pp; Japanese.
 XX
 CC The invention relates to a novel purified protein having Elki
 CC phosphorylation activity and/or an activity of activating Elki
 CC immunomodulator, virucide, cyostatic, anti-allergic, antineumatic,
 CC antidiabetic, antidiabetic, antidiabetic, and anti-HIV activity. The
 CC polynucleotides may have a use in gene therapy. The gene and its encoded

CC The invention provides polypeptides capable of stimulating nuclear factor (NF)-kappaB-dependent transcription or p38-dependent transcription, referred as pellino polypeptides. The pellino polypeptides are useful for identifying modulators that alter the pellino polypeptide and pellino dominant-negative activity. They are also useful for identifying compounds that inhibit the binding activity of the polypeptides and to study cell-signal transduction. They are useful for preventing or treating infection by a pathogen such as virus, bacterial, fungi, algae or protozoa, or inhibiting apoptosis. Dominant-negative pellino polypeptides are useful for treating inflammatory conditions such as asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and also for inhibiting mitogen activated protein (MAP) kinase-activated pathways. Pellino polypeptides and polymucosides are useful to identify small molecule inhibitors of protein association or function of Pellino, and other molecules involved in interleukin (IL)-1 signaling. The present sequence represents a human pellino-1 polypeptide coding sequence

Alignment Scores:

Pred. No.: 9 05e-166 Length: 1257
Score: 1917.00 Matches: 342
Percent Similarity: 90.5% Conservative: 38
Best Local Similarity: 81.4% Mismatches: 38
Query Match: 83.7% Indels: 2
Gaps: 1

US-10-041-030-4 (1-420) x ABL58447 (1-1257)

QY 1 MetPheSerProGlyInGluGluHisCysAlaProAsnIysGluProValIysTyrGly 20
DB 1 ATGTTTCTCTGATCAAGAAATCAT-----CCATCTAAGACGCGTAATAATAGT 54
QY 21 GluLeuValIleuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGly 40
DB 55 GAACCTATGCTTGAATGATTAATGATCTCTCCCAACGCTGATAGAGAAAGGAGAA 114
QY 41 SerArgPheAlaLeuTyrIysArgProIysAlaAsnGlyValIysPProSerThrValHis 60
DB 115 AGTAGGTTCTTGTGTTAAAGACCTAAGGCAATGAGGTGAAGGCCAGACGTCAT 174
QY 61 ValIleSerThrProGlnAlaSerIysAlaIleSerCysIysGlyGlnHisSerIleSer 80
DB 175 ATTGCTTGTACTCTCTCGGCGCCAGACCTGTGTGTAATATCTCATGACACACCGAT 234
QY 81 TyrThrIleuSerArgAsnGlnThrValValIleGlyTyrThrHisAspIysAspThrAsp 100
DB 235 TATATCTTATCTCGGCGCCAGACCTGTGTGTAATATCTCATGACACACCGAT 294
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
DB 295 ATGTTTCAGATTGGCCGCTGACAGAAAGCCCAATATTTGTAGTACGACACCGTT 354
QY 121 SerGlySerIleAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 355 CCGGAAGTCAAGTATCTCGATACACAGTCACTACAAAGCACTATATCAAGTTGGC 414
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
DB 415 TGCGAATCATATGTGAACCGAATCCCTCTTACACACGAGATTTTGGCGAGGCTTT 474
QY 161 ArgSerSerIysAsnIlePheIleuGlyGluIysValIleAlaIysThrIysAsnProAsnGly 180
DB 475 GACTCATCAAAAACATCTTTTGGGAGAGAGGCTCCCAATGGAAGACATCAGATGGA 534
QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValIleHisPProArgGlyGlyPheThr 200
DB 535 CAGATGATGGCTTGACACCTAATGCTGTTCTTGTGATGATCCACCGAATGGTTACA 594
QY 201 GluGlnSerGlnProGlyValITrpArgGluIleSerValCysGlyAspValTyrThrLeu 220
DB 595 GAAGACTCCAAAGCTGGAATATGAGAGAAATATCGGTGTGTGGAAATGATTTAGCTTA 654

QY 221 ArgGluThrArgSerAlaGlnGlnAlaArgGlyIleLeuValGluSerGluThrAsnValLeu 240
DB 655 CGTGAACCAAGATCGGCTGACGAGAGAGGAAATAAGTGGAATTAACCAATCATGTTA 714
QY 241 GlnAspGlySerIleuIleAspLeuCysGlyAlaThrLeuLeuThrAlaAspGly 260
DB 715 CAAGATGGCTCGTTAATTGACCTCTGTGGCAACATTTGTATGGGCTACGCAAGGCG 774
QY 261 LeuPheHisThrProThrGlnIysHisIleGluAlaLeuArgGlnGluIleAsnAlaIle 280
DB 775 CTTCACCACTCTCAACCGTAAGCAATTAAGCTTTAAGACAGAAATCAATGACGCA 834
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgIysGlu 300
DB 835 CGACCTCAGTGGCCCTGATGGTTCAACACACTGATCTTCTATGTAAGAGGAAGAC 864
QY 301 ValValGluGluIysGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
DB 895 GTTGATATGAAAAACACCATGAGTATATCTAACTGCGCGCATGTATCATGGCTATCAT 954
QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluAlaArgGluCysProMetCysArgThr 340
DB 955 AACTGGGAGAAACAAAGAAACGTGATGGCAAGATCGTAATGTCCTATGTAGAGTCT 1014
QY 341 ValGlyProTyrValProLeuTyrLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
DB 1015 GTTGTCCTCATTTCTCTGTGGCTTGATGAGAGCTGATTTATATGACGCGCGC 1074
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluIysSerAlaIysTyr 380
DB 1075 CTTCAACCATGCTGTTAGCCGCTGTGGGCAATGTGTGTTCAAGAAAGCAATGCTAT 1134
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIleAspPProPheCys 400
DB 1135 TGGTCCAGATCCACCTCTCTCATGGTACTCAATCTTTCATGACACGCTGCTTTGT 1194
QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
DB 1195 GCACATCATGTGGCTGTGTAACAAAGCTATCATGACCTATTTTTCAGAGACCTTAGAC 1254

RESULT 11
ADP48666 standard; cDNA, 1257 BP.
ID ADP48666;
AC ADP48666;
XX 09-SEP-2004 (first entry)
DT XX
DE Human Pellino-1 encoding cDNA SEQ ID NO:3.
XX
KW nuclear factor kappa B dependent transcription inhibitor;
KW NF-kB-dependent transcription inhibitor;
KW p38-dependent transcription inhibitor; Pellino-1;
KW interleukin 1 receptor-associated kinase 4;
KW IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
KW antiarthritic; antidiabetic; antiinflammatory; antirheumatic; antidiabetic;
KW gastrointestinal; neuroprotective; nootropic; IL-1 antagonist;
KW IRAK protein kinase family inhibitor;
KW NF kappa B inducing kinase inhibitor; asthma; rheumatoid arthritis;
KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KW atherosclerosis; Alzheimer's disease; human; chromosome 2; gene; ss.

OS Homo sapiens.
XX
FH Key 1.1257 Location/Qualifiers
FT CDS /tag= a
FT FT /product= "pellino-1"
XX
PN NC02004053092-A2.
XX
PD 24-JUN-2004.

XX ABX05095;
AC
XX 17-JAN-2003 (first entry)
DT
XX Human novel polynucleotide #110.
DE
XX Human; gene; ser; genetic disorder; gene mapping; medical imaging; cancer;
XX neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
XX Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
XX osteoarthritis; periodontal disease; liver fibrosis; viral infection;
XX fungal infection; bacterial infection; autoimmune disease; diabetes;
XX atopic dermatitis.
OS
XX Homo sapiens.
XX MO200274961-A1.
XX
XX 26-SEP-2002.
PD
XX 14-MAR-2002; 2002WO-US005109.
PF
XX 15-MAR-2001; 2001US-00810173.
PR
XX (HYSR-) HYSEQ INC.
XX
XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Weinman T, Wang J, Wang D, Drmanac RT;
XX
XX WPI; 2003-040556/03.
DR P-SDB; ABU00017.
XX
XX New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.
XX
XX Claim 1; SEQ ID NO 110; 235pp; English.
PS
XX The invention relates to human polynucleotides and the polypeptides they
XX encode. The polynucleotides and polypeptides are useful in diagnostics,
XX forensic, gene mapping, medical imaging, identification of mutations,
XX responsible for genetic disorders or other traits, assessing biodiversity
XX and producing many other types of data and products dependent on DNA and
XX amino acid sequences. They are also useful for preventing, treating or
XX ameliorating medical conditions, such as cancer, neurodegenerative
XX disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
XX disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
XX periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
XX bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
XX Sequences ABX04986-ABX05511 represent human polynucleotides of the
XX invention. Note: The sequence data for this patent is not represented in
XX the printed specification but is based on sequence information supplied
XX by the European Patent Office
XX
XX Sequence 1304 BP; 385 A; 270 C; 305 G; 344 T; 0 U; 0 Other;
SO
XX
XX Alignment Scores:
XX Pred. No.: 9.52e-166 Length: 1304
XX Score: 1917.00 Matches: 342
XX Percent Similarity: 90.5% Conservative: 38
XX Best Local Similarity: 81.4% Mismatches: 38
XX Query Match: 83.7% Indels: 2
XX DB: 8 Gaps: 1
XX
XX US-10-041-030-4 (1-420) x ABX05095 (1-1304)
XX
XX 1 MetPheSerProGluGlnGluGlnHisCysAlaProAlaSerGluProValIleGlyTyrGly 20
DB 1 ATGTTTCTCTCTGATCAAGAAATCAT-----CCATCTAAAGCACAAGAAATTAATGCT 54
XX 21 GluLeuValIleGluGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGlyLeu 40

DB 55 GAACATCTGCTTTAGGATATATATGGCTCTCCCAATATGCGCATAGAGAAAGAGAA 114
XX
XX 41 SerArgPheAlaLeuYrlyAsArgProLyAlaAsnGlyValIlyProSerThrValHis 60
DB 115 AGTAGGTTGCTTGTATTAAGACCTTAAGCAAAATGGGGTGAAGCCGACACTGTGCAT 174
XX
XX 61 ValIleSerThrProGlnAlaSerLyAlaIleSerCysIleGlyGlnHisSerIleSer 80
DB 175 ATTCCTTGAATCTCTCAGGCTGCAAGGCAATAGCAACAAAGACAGATATCATATCA 234
XX
XX 81 TyrThrLeuSerArgAsnGlnThrValIleGlyTyrThrHisAspIlyAspThrAsp 100
DB 235 TATACCTTGTCTCGGCGCCAGACTGTGTGGTTGAATATATCATATGACAGCAACAGAT 294
XX
XX 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
DB 295 ATGTTTCAGATTGGCCGGTGTGACTGAAGCCCATTTGATTTGTATGACTGACAGGTT 354
XX
XX 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 355 CCGGAAGTCAAGTAAATTTCTGATACAGTCACTGTCMAAGCATATATCATAGATTGGC 414
XX
XX 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
DB 415 TGCAGATCATATGTGAGACGAATCCTCCCTTACAGACGGAATTTATGCTGCAGGATTT 474
XX
XX 161 AspSerSerLyAsnIlePheLeuGlyGluValAlaIleThrIlyAsnProAspGly 180
DB 475 GACTCATCAAAAACATCTTCTTGGGAGAAAGGCTGCMAATGAGAAAGACATCATAGTGA 534
XX
XX 181 HisMetAspGlyLeuThrThrAsnGlyValIleValMetHisAspArgGlyGlyPheThr 200
DB 535 CAGATGATGCGCTTGAACCACTAATGCTTCTGTATGATGATCAAGCAATGGCTTCACA 594
XX
XX 201 GluGluSerGlnProGlyValIleTyrArgGlyIleSerValCysGlyAspValIleThrLeu 220
DB 595 GAAGACTCCAGCGCTGGAATATGAGAGAAATATCGGTGTGGAATGATTAATTAACCTTA 654
XX
XX 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGlnThrAsnValLeu 240
DB 655 CGTGAACCGCATGCGCTCGCAGAGAGAAATATGATGAAATTAACCAATCACTTA 714
XX
XX 241 GluAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
DB 715 CAAAGTGGCTCGTAAATTAATGACTGTGTGCGCAACATTTGATGGGCTACAGAAAGC 774
XX
XX 261 LeuPheHisThrProThrGlnIlyHisIleGluAlaLeuArgGlnIleAsnAlaIle 280
DB 775 CTTTCCACACTCCTACCGCGAAGCATTTTGAAGCTTTAAGACAGAAATCAATGCGACA 834
XX
XX 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnAlaGlyLeu 300
DB 835 CGACCTCAGAGCCCTGTAGGGTTCAACACATGCAATTTCTCTGATGAAAGAGAAAGAAC 894
XX
XX 301 ValIleGluGlnIlyAsnGlnProTyrAlaTyrLeuSerCysGlyHisIleGlyTyrHis 320
DB 895 GTTGATGATAAACCAACCATGCGGTATCTTAACCTGCGGCCATGTACATGCGCTATCAT 954
XX
XX 321 AsnTyrGlyHisArgSerAspThrGlnAlaAsnGlnArgGlyCysArgProMetCysArgThr 340
DB 955 AACTGGGAAACAAAGAAAGAAAGCATGATGAAAGATCGTAAATGCTCTATGTAGGTTCT 1014
XX
XX 341 ValGlyProTyrValProLeuThrPleuGlyCysGlnAlaGlyIleThrValAlaPheAlaGly 360
DB 1015 GTTGTGCTCCATATGCTCTGTGTGGCTTGAATGGAAGCTGATTTATATGACGCGGCG 1074
XX
XX 361 ProProThrHisAlaPheThrProCysGlyHisIleValCysSerGlyIlySerAlaIlyTyr 380
DB 1075 CTTCCAAACCATGCGTTTACGCCGTGTGGCATGTGTGTTTCAAAAAGACAATGCGCTAT 1134
XX
XX 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIleAspCysArgProPheCys 400

DB 1135 TGGTCCAGATCCCACTTCATGTAATCACTTTTCATGACGCTGTCCTTTGT 1194
 QY 401 AlAThGlnIeuValGlyGlnIAsnCylIeIysIeuLeuPheGlnGlyProIleAsp 420
 DB 1195 GCACATCAATGTTGGCTGGTGAACAGGCTACATCACTATTATTATTCAAGACCTCTAGAC 1254

RESULT 13
 AAL46320
 ID AAL46320 standard; cDNA; 3222 BP.
 AC AAL46320;
 DT 19-JUL-2002 (first entry)
 XX Human M30 variant D coding sequence SEQ ID NO: 9.
 DE
 XX Neurodegenerative disease; M30; M31; M32; M33; stroke;
 KW fragile X syndrome; Huntington's disease; Parkinson's disease;
 KW Alzheimer's disease; multiple sclerosis; ovarian cancer;
 KW neurodegeneration; immune disorder; autoimmune disease; allergy;
 KW infection; leukaemia; inflammation; neuroprotective; cerebroprotective;
 KW immunosuppressive; cytostatic; nootropic; antiparkinsonian; antiallergic;
 KW vaccine; antiinflammatory; gene; ss.
 XX Homo sapiens.
 OS
 XX NO200221138-A2.
 PN
 XX 14-MAR-2002.
 PD
 XX 07-SEP-2001; 2001WO-BP010366.
 PF
 XX 07-SEP-2000; 2000US-00657479.
 PR
 XX (AXAR-) AXARON BIOSCIENCE AG.
 PA
 PI Schneider A, Hienisch H, Rosner M, Klugmann M, Naim J;
 PI Eisenhardt G, Kuner R, Lanahan A, Worley P, Spielvogel D, Scheek S;
 PI WPI; 2002-292287/33.
 DR P-PSDB; AAO17519.
 DR
 XX Diagnosis of neurodegenerative disease comprises detecting level of M30-
 PT family proteins.
 PT
 XX Claim 3; Page 97-101; 130pp; German.
 PS
 XX The present invention relates to a method of diagnosing neurodegenerative
 CC diseases, comprising determining the concentration of a protein in a body
 CC sample, where the protein may be M30 or a variant thereof, M31, M32 or
 CC M33. The method is used to diagnose neurodegenerative diseases,
 CC particularly stroke but also e.g. fragile X syndrome, Huntington's,
 CC Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
 CC overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
 CC especially ovarian cancer. The proteins can be used to identify specific
 CC ligands, potentially useful for treating neurodegeneration, immune-system
 CC disorders (e.g. autoimmune diseases, allergy, viral infection, leukaemia,
 CC inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
 CC between the proteins and the protein kinase IRAK-1 can be used to treat
 CC neurodegeneration. The present sequence is a coding sequence of a protein
 CC used in the method of the invention

XX
 SQ Sequence 3222 BP; 980 A; 576 C; 641 G; 1025 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3,27e-165 Length: 3222
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.5% Conservative: 38
 Best Local Similarity: 81.4% Mismatches: 38
 Query Match: 83.7% Indels: 2
 DB: Gaps: 1

US-10-041-030-4 (1-420) x AAL46320 (1-3222)

QY 1 MetPheSerProGlyGlnGlnIAsnCylIeIysIeuLeuPheGlnGlyProIleAsp 20
 DB 104 ATGTTTCTCTCATCAAGAAATTCAT-----CACTTAAACACACGATTAATATGCT 157
 QY 21 GlnIeuValIleuGlyIleuValIleuValIleuValIleuValIleuValIleuVal 40
 DB 158 GAATCATTTCTCTTAGGGTATATGGGTCTCTCCAAATGGCATATAGAGAGAGAGAAA 217
 QY 41 SerArgPheAlaIeuTyrIleuValIleuValIleuValIleuValIleuValIleuVal 60
 DB 218 AGTAGGTTTCTCTTGTAAAGACCTTAAGCAAAATGGGTGAAGCCACACCTGTCAT 277
 QY 61 ValIleSerThrProGlnIleSerIleValIleuValIleuValIleuValIleuValIleuVal 80
 DB 278 ATGCTTGTATCTCTCAAGCTGCAAGGCAATAGCAACAAACACGATGATATCA 337
 QY 81 TyrThrIleuSerArgAngIleuValIleuValIleuValIleuValIleuValIleuVal 100
 DB 338 TATACCTTGTCTGGGCGCCGACCTGTGTGTGATATATCTCATGACACACACACAT 397
 QY 101 MetPheGlnValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 120
 DB 398 ATGTTTCAGATGGCCGCTGACCTGAAGCCCATTTGATGATGATGATGATGATGATGAT 457
 QY 121 SerGlySerGlnIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 140
 DB 458 CTTGAAGTCAAGATTCATCTGATACACATGCAATCAAGCACTATATCAAGATTTGCC 517
 QY 141 CysArgIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 160
 DB 518 TGCAGATATCATATGATGACGATCTCTCTTACACACGATTTATGCTGACGATTT 577
 QY 161 AspSerSerIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 180
 DB 578 GACTCATCAAAAACATCTTTCTGGGAGAAAGCTGCCAATGAAAGACATGACGATGGA 637
 QY 181 HisMetArgPheIleuThrIleuValIleuValIleuValIleuValIleuValIleuVal 200
 DB 638 CAGATGATGGCTGACCACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 697
 QY 201 GlnGluSerGlnProGlyValIleuValIleuValIleuValIleuValIleuValIleuVal 220
 DB 698 GAAGACTCCAGGCTGGAATATGAGAAATATGAGTGTGTAATGATTTAGCTTA 757
 QY 221 ArgIleuThrArgSerArgIleuValIleuValIleuValIleuValIleuValIleuVal 240
 DB 758 CGTGAACCAAGATCGGCTGACGAGAGAGAAATGATGAAATGAAATCAATCACTTA 817
 QY 241 GlnAspGlySerIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 260
 DB 818 CAAGATGGCTCGTAAATGACCTCTGTGGCAACATTTGATGGGTATCGACGAGAGGC 877
 QY 261 LeuPheHisThrProThrGlnIleuValIleuValIleuValIleuValIleuValIleuVal 280
 DB 878 CTTTCCACACTCTCAACCGAAGCATTTTAAAGCTTTTAAAGCAAGATCAATGACGA 937
 QY 281 ArgProGlnCysProValGlyIleuValIleuValIleuValIleuValIleuValIleuVal 300
 DB 938 CGACTCAGGCTCTGATAGGTTCAACACACTGACATTTCTGATGATGAAAGAGAAAGAC 997
 QY 301 ValValIleuValIleuValIleuValIleuValIleuValIleuValIleuValIleuVal 320
 DB 998 GTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1057
 QY 321 AsnThrGlyHisArgSerArgThrGlnIleuValIleuValIleuValIleuValIleuVal 340
 DB 1058 AACTGGGGAACCAAGAGACCTGATGAAAGATCGAATGCTTATGTTAGTCT 1117
 QY 341 ValGlyProTyrValProIleuValIleuValIleuValIleuValIleuValIleuVal 360
 DB 1118 GTTGTCCCATTTCTCTGT 1177

ID	Accession	Standard	CDNA	3526 BP
QY	361	ProProthThiSaIalphehrProCySgIyHsIvalCySesgIyUySerIaySTyr	380	
Db	1178	CTCCCAACCAACGCGTTTACCGCGGTGCGCATGTGTCTTCAGAAAGACAACTCCGCTT	1233	
QY	381	TrpSerGlnIleProLeuProHisgIyThrHsIalAphesIalAalAaCySFrProhCyS	400	
Db	1238	TGGTCCCAATCCCACTTCTCCTCATGGTACTCATTACTTTTCATGACAGCGTCCCTTGGT	1297	
QY	401	AlaThrGlnLeuValGlyGluGlnAnCySIIeIySleuIIlePheGlnGlyProIleAap	420	
Db	1298	GCACATCACTGGCTGGTGTAAACAGGCTACATCAACGACTTATTTTTCAGAGACCTCTAGAC	1357	

RESULT 14
AAL46319

ID	Accession	Standard	CDNA	3526 BP
AC	AAL46319			
XX	19-JUL-2002	(first entry)		
DE	Human M30 variant C coding sequence SEQ ID NO: 7.			
XX				
KW	Neurodegenerative disease; M30; M31; M32; M33; stroke;			
KW	fragile X syndrome; Huntington's disease; Parkinson's disease;			
KW	Alzheimer's disease; multiple sclerosis; ovarian cancer;			
KW	neurodegeneration; immune disorder; autoimmune disease; allergy;			
KW	infection; leukaemia; inflammation; neuroprotective; cerebroprotective;			
KW	immunosuppressive; cytostatic; nocotropic; antiparkinsonian; antiallergic;			
KW	viaticid; antinflammatory; gene; ss.			
OS	Homo sapiens.			
XX				
PN	WO200221138-A2.			
PD	14-MAR-2002.			
PF	07-SEP-2001; 2001MO-EP010366.			
XX				
XX	07-SEP-2000; 2000US-00657479.			
XX				
PA	(AXAR-) AXARON BIOSCIENCE AG.			
PI	Schneider A, Hienisch H, Rosener M, Klugmann M, Naim J,			
PI	Eisenhardt G, Kuner R, Lanahan A, Worley P, Spielvogel D, Scheek S;			
XX				
DR	WPI; 2002-292287/33.			
XX				
PT	P-PSDB; AAO17518.			
XX				
XX	Diagnosis of neurodegenerative disease comprises detecting level of M30-			
XX	family proteins.			
PS	Claim 3; Page 91-95; 130pp; German.			
XX				
CC	The present invention relates to a method of diagnosing neurodegenerative			
CC	diseases, comprising determining the concentration of a protein in a body			
CC	sample, where the protein may be M30 or a variant thereof, M31, M32 or			
CC	M33. The method is used to diagnose neurodegenerative diseases,			
CC	particularly stroke but also e.g. fragile X syndrome, Huntington's,			
CC	Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also			
CC	overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,			
CC	especially ovarian cancer. The proteins can be used to identify specific			
CC	ligands, potentially useful for treating neurodegeneration, immune-system			
CC	disorders (e.g. autoimmune diseases, allergy, viral infection, leukaemia,			
CC	inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction			
CC	between the proteins and the protein kinase IRAK-1 can be used to treat			
CC	neurodegeneration. The present sequence is a coding sequence of a protein			
XX	used in the method of the invention			
SQ	Sequence 3526 BP; 1076 A; 618 C; 685 G; 1147 T; 0 U; 0 Other;			

Alignment Scores:

Pred. No.:	3,7e-165	Length:	3526
Score:	1917.00	Matches:	342

Percent Similarity:	90.5%	Conservative:	38
Best Local Similarity:	81.4%	Mismatches:	3
Query Match:	83.7%	Indels:	2
DB:	6	Gaps:	1
US-10-041-030-4 (1-420) x AML46319 (1-3526)			
QY	1 MetPheSerProGlyGlnGluGluHhCYsaAlaProAenlySGuIProValIysTYrGly	20	
DB	408 ATGTTTCTCTGATTCAGAAATATCAT-----CATCTTAAGCACCAGTAAATATAGGT	461	
QY	21 GluLeuValValLeuGlyTYrAsnGlyAlaLeuProAenGlyAspArgGlyArgArgLys	40	
DB	462 GAATCTCATTTGCTTAGGGTATAATAGGGCTCTCCCAATGGCGATAGAGAAAGAGAA	521	
QY	41 SerArgPheAlaLeuTYrLYsArgProLYsAlaAenGlyValLYsProSerThValHis	60	
DB	522 AGTGGTTTGCTTTGTTTAAAGACCTTAAGCAATGGGGTGAAGCCGACGACGTGCAT	581	
QY	61 ValIleSerThrProGlnAlaSerLYsAlaAlaIleSerCyLYsGlyGlnHisSerLleSer	80	
DB	582 ATTCCTGTACTCTCTCAGGCTGCAAGGCAATTAAGCAACAAAGCCAGCATAGCATATCA	641	
QY	81 TYrThrLeuSerArgAsnGlnThrValValValGlyTYrThrHisAspLYsAspThrAsp	100	
DB	642 TATATCTTTGCTCGGGCCGACAGCTGTGGTTGAAATTAATCTCATGACAGAACACAGAT	701	
QY	101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle	120	
DB	702 ATGTTTACGATGGCCGGGTGACATGAAGCCCATGATGATTTTGTATGTAATGACACGGGT	761	
QY	121 SerGlySerGlnAenThrAspGluAlaGlnIleThrGlnSerThrLleSerArgPheAla	140	
DB	762 CCTGAATCAACAAATATCTTGATACACACAGTCAGTACAAAGCATTATATCAAGATTTGCC	821	
QY	141 CysArgIleValCYsAspArgAsnGlnProTYrThrAlaArgIlePheAlaAlaGlyPhe	160	
DB	822 TCGAATATCATATGTGAACGGAATCTCTCCCTTACAGACGGAATTTATGCTGCGAGATTT	881	
QY	161 AspSerSerLYsAsnIlePheLeuGlyGlyLYsAlaAlaLYsTYrLYsAsnProAspGly	180	
DB	882 GATCATCATAAACAAATCTTTCTTGCGGAGAAAGGCTGCCAAATGGAAGACATCAGATGA	941	
QY	181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr	200	
DB	942 CAGATGATGGCTTGACCACTAAAGGTGTTCTTGATGTCATCCAGCAAGGATGCTTCA	1001	
QY	201 GluGluSerGlnProGlyValTYrArgGlnLleSerValCYsGlyAspValTYrThrLeu	220	
DB	1002 GAAAGCTCCAAAGCCCTGGAAATGGAAGAAATATCGGTGTGGAAATGATATTAAGCTTA	1061	
QY	221 ArgGluThrArgSerAlaGlnGlnAlaArgGlyLYsLeuValGluSerGluThrAsnValLeu	240	
DB	1062 CGTGAACACCAATGGGCTCAGACAGAGAGGAAATGGCGAAATTAAGAACCAATCAGTTA	1121	
QY	241 GlnAspGlySerLeuIleAspLeuCYsGlyAlaThrLeuLeuTYrArgTYrAlaAspGly	260	
DB	1122 CAAGATGGCTCGTAAATGACCTCTGTGGTCAACATTTGATGGCGTACGCAAGAGGC	1181	
QY	261 LeuPheHisThrProThrGlnLYsHisIleLeuAlaLeuAArgGlnGlnIleAsnAlaIa	280	
DB	1182 CTTTCCCAACATCTCTGACCGTAAACATTTTAACACTTTTAAGACAGAAATCAATGACGA	1241	
QY	281 ArgProGlnCYsProValGlyLeuAenThrLeuAlaPheProSerLleAsnArgLYsGlu	300	
DB	1242 CGACCTCATGCTCCCTGTAGGGTTCACACACTTACATTTCTTACTATATGAAGAAAGAC	1301	
QY	301 ValValGluGluLYsGlnProTYrAlaTYrLeuSerCYsGlyHisValHisGlyTYrHis	320	
DB	1302 GTTGTAGATGAAAAACAACATGGGTATATCTTAATATGCGCGCATGTATCATGGGTATCAT	1361	
QY	321 AsnTYrGlyHisArgSerAspThrGluAlaAenGluAArgGluCYsProMetCYsArgTYr	340	

Db 1362 AACTGGGAAACAAAGAGACGATGAGAAAGATGATGCTATGTAGAGTCT 1421
 QY 341 ValGlyProTyrValProLeuTyrPLeuGlyCysGluAaGlyPheTyrValAlaPalaGly 360
 Db 1422 GTTGTCCTCTATGCTCTGCTGCTGATGAGATGATGATTTATGTGACGCGGC 1481
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluYsSerAlaLysTyr 380
 Db 1482 CCGCAACCAACGATGCTTACCGCTGTGGCATGTGTTCACAAAAGACACTGCTAT 1541
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaCysProPheCys 400
 Db 1542 TGTGTCACAGATCCACTTCTCTCATGTACTCATCTTTATCATGACGCTGCTTTGT 1601
 QY 401 AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
 Db 1602 GCACATCAGTTGCTGTGTGAACAAGCTACATCAGACTTATTTTCAAGACCTTAGAC 1661
 RESULT 15
 ID ADF83102 standard; cDNA; 7136 BP.
 AC ADF83102;
 XX 26-FEB-2004 (first entry)
 DT
 XX Human Pellino homologue 1 gene, overexpressed in cancer.
 DE Human Pellino homologue 1; Pellinot; cancer; cytosolic; vaccine;
 KM Gene therapy; gene; oncogene; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH CDS 4039..5295
 FT /*tag= a
 XX MO2003100000-A2.
 PN
 XX 04-DEC-2003.
 PD
 XX 22-MAY-2003; 2003MO-US016049.
 PF
 XX 24-MAY-2002; 2002US-0382606P.
 PR 25-JUL-2002; 2002US-0398099P.
 XX
 PA (TULA-) TULARIK INC.
 XX
 PI Li J, Mu D, Yang J;
 XX
 XX WPI; 2004-035118/03.
 DR P-PSDB; ADF83103.
 DR GENBANK; NM_020651.
 XX
 PT Diagnosing a cancer in a mammal comprises determining RecQL5, CTXL,
 PT USP1, MCL1, or Pellino 1 gene copy number in a biological sample from a
 PT region of the mammal that is suspected to be precancerous or cancerous.
 XX
 PS Claim 179; SEQ ID NO 9; 174bp; English.
 XX
 CC The present sequence is that of the gene encoding human Pellino homologue
 CC 1 (Pellinot), a protein associated with the kinase domain of activated
 CC Pelle. The invention is based on the finding of the overexpression of
 CC Pellinot and other genes (RecQL5, CTXL, USP13 and MCL1) in certain
 CC cancers, including breast cancer, colon cancer, lung cancer and ovarian
 CC cancer, and the frequent amplification of these genes in cancer cells.
 CC The genes, and their expression products, can be used diagnostically or
 CC as targets for cancer therapy. They can also be used to identify and
 CC design compounds useful in the diagnosis, prevention and therapy of
 CC tumours and cancers, in vaccine development, and in methods for
 CC determining the efficacy of a treatment regime. A claimed method for
 CC inhibiting cancer or precancerous growth, especially in ovarian or lung
 CC tissue, uses an inhibitor that interacts with Pellinot DNA or RNA. The

CC inhibitor is a small interfering RNA (siRNA), microRNA (miRNA), an
 CC antisense RNA, and antisense DNA, a decoy molecule, a decoy DNA, a
 CC ribozyme or small molecule.
 XX
 SQ Sequence 7136 BP; 1945 A; 1333 C; 1399 G; 2459 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 9 67e-165 Length: 7136
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.5% Conservative: 38
 Best Local Similarity: 81.4% Mismatches: 38
 Query Match: 83.7% Indels: 2
 DB: Gaps: 1
 US-10-041-030-4 (1-420) x ADF83102 (1-7136)
 QY 1 MetPheSerProGlyGlnGluHisCysAlaProAsnLysGluProValLysTyrGly 20
 Db 4039 ATGTTTCTCTGATCAAGAAATCAT-----CAATCAAAGCACAGTAAATATGTT 4092
 QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGly 40
 Db 4093 GAACATATGCTTAAAGCTATATAGGCTCTCCCAATAGCGATAGAGAGAGAA 4152
 QY 41 SerArgPheAlaLeuTyrTyrAspArgProLysAlaAsnGlyValLysProSerThrValHis 60
 Db 4153 AGTAGGTTTGCTTTGTTAAAGACTTAAGCAAAAGGGGTGAAGCCAGCATGTGCAT 4212
 QY 61 ValIleSerThrProGlnLaseLysLysAlaIleSerCysLysGlyCysHisSerIleSer 80
 Db 4213 ATTGCTGTACTCTCAGGCTCGAAGGCAATAGACAAAGAACACAGCATGACATCTCA 4272
 QY 81 TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAspLysAspThrAsp 100
 Db 4273 TAACTTGTCTCGGGCCAGACTGTGGTGAATATATCATATACATGACAGACAGAT 4332
 QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
 Db 4333 ATGTTTCAGATTGGCCGGTGCAGCTGAAGCCCCCATGATTGTGTAACTGACAGCGTT 4392
 QY 121 SerGlySerGlnAsnThrAspGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 4393 CCGAAGATCAAGATTAATCTGATACAGTCAAGTCAAGACATATACAGATTGGC 4452
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
 Db 4453 TGCAGATCATATGTGACGAGATCTCTTATACAGCAGGATTTATCTGACGATTT 4512
 QY 161 AspSerSerLysAsnIlePheLeuGlyGlyValAlaLysIleThrLysAsnProAspGly 180
 Db 4513 GACTCATCAAAAACATCTTTCTGGGAGAAAGGCTGCCAAATGGAAGCATCAGATGGA 4572
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgLysGlyPheThr 200
 Db 4573 CAGATGATGGCTTGACCACTAATGCTGTCTTGATCATATCCAGCAATGGGTTCAACA 4632
 QY 201 GluGluSerGlnProGlyValTyrPargGluIleSerValCysGlyAspValTyrThrLeu 220
 Db 4633 GAAAGATCCAGGCTGATGATATGAGAGAAATATGCTGTGTGGAATATATTACCTTA 4692
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
 Db 4693 CGGAAACCAAGATCGGCTCAGACAGAGCAAAATGTGTGAATGGAACCAATCAGTTA 4752
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
 Db 4753 CAAGATGCTCGTAAATTAACCTCTGTGTGCAACTTTTAAGGCTGACAGAGGC 4812
 QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAla 280
 Db 4813 CTTTCCACATCTTACCGTGAAGCATTTTGAAGAGCTTTAAGACAGGAATCATATGACGA 4872
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300

Db	4873	CGACCTCAGTCCCTTGAGGGTTCAACACCTTACGATTTCTTAGTAGAGAGAAAGAC	4932
Qy	301	ValValGluGluIuyluysGlnProTrrPalaTyLeuSerCySglYhiValhIagLYrHis	320
Db	4933	GTTTAGTAGTAAAAACAACATGGGGTATATTAAACGCGGCATGTACATGGTATCAT	4992
Qy	321	AsnTrpGlyhiIarGserAspThrGluAlaangIuargIuCyPrometCyasrGthr	340
Db	4993	AACGGGGAAACAAAGAAAGACGGATGGAAAGATGGTAATGGTATGTAGAGTCT	5052
Qy	341	ValGlyProTyValProLeuTrpLeuGlyCySgluAlaGlyPheTyValAspAlaGly	360
Db	5053	GTGTGTCCTTATGTTCCCTCTGTGGCTTGGATGTAAGCTGGATTTTATGTGACGCGGC	5112
Qy	361	ProProthrhIsaIaPheThrProCySglYhiIstValCySserGluIuysSerAlaIuTyTr	380
Db	5113	CCTCCAAACCCATGGGTGATGACCGGTGGGCGATGTGTGTAGAAAGACAACGTCCAT	5172
Qy	381	TrpSerGlnIleProLeuProhIstGlyThrHisIaIaPhehIsaIaIaCyProPhCyS	400
Db	5173	TGGTCCCAAGATCCCACTTCATAGTAGTATATCTTTCAAGAGCGCTGCTCTTTGT	5232
Qy	401	AlaThrGlnIuValGlyGluIuAsnCyIleIysLeuIlePheGlnGlyProIleAsp	420
Db	5233	GCAATCATGTTGGCTGGTGAACAAAGCGTACATCACTTATTTTTCANAGACCTCTAGAC	5292

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocollaboration Ltd.

OM protein - nucleic search, **using frame p2n model**

Run on: March 2, 2006, 05:01:06 ; Search time 6947 Seconds
(without alignments)
3436.627 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MESPGEHHCAPKPKVYKX.....ATQLVGEQNCIKLIPGSPID 420

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delcxt 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSMB.spool/US10041030/runat_01032006_134404_21821/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss808
-USER=US10041030.CCN_1_1_4939 @runat_01032006_134404_21821 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb ba: *
2: gb in: *
3: gb env: *
4: gb om: *
5: gb ov: *
6: gb pat: *
7: gb ph: *
8: gb pr: *
9: gb ro: *
10: gb sts: *
11: gb sy: *
12: gb un: *
13: gb vi: *
14: gb hcg: *
15: gb pl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match Length	DB ID	Description
1	2290	100.0	1565 6	AX817203 Sequence
2	2290	100.0	1811 8	BC009476 Homo sapi
3	2290	100.0	5579 6	BD190328 Elkt phos

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	CDS
AX817203	Sequence 15 from Patent WO0221138.	AX817203	AX817203	GI:39722594			Homo sapiens (human)	Schneider, A., Hiemisch, H., Rosener, M., Klugmann, M., Naim, J., Eisenhardt, G., Kuner, R., Lananan, A., Worley, P., Spielvogel, D. and Scheek, S.	The m30 gene family and the utilization thereof	Patent: WO 0221138-A 15 14-MAR-2002;	Axaron Bioscience AG (DE)	Location/Qualifiers 1. 1565 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 51..1313 /note="unamed protein product; codierende Bereich der M3 CDNA des Menschen"	

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	CDS
AX817203	Sequence 15 from Patent WO0221138.	AX817203	AX817203	GI:39722594			Homo sapiens (human)	Schneider, A., Hiemisch, H., Rosener, M., Klugmann, M., Naim, J., Eisenhardt, G., Kuner, R., Lananan, A., Worley, P., Spielvogel, D. and Scheek, S.	The m30 gene family and the utilization thereof	Patent: WO 0221138-A 15 14-MAR-2002;	Axaron Bioscience AG (DE)	Location/Qualifiers 1. 1565 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 51..1313 /note="unamed protein product; codierende Bereich der M3 CDNA des Menschen"	

/codon_start=1
/protein_id="CAE84330.1"
/db_xref="GI:39722595"
/translation="MPSPGQBEHCAPKPEVKKYKGLVYLGNALPNCGRKSRPA
LYRPRANGVKPSTVHTVISTPQASKAISCGQHSISTTISRACKIVCDRNPYTAI.PAAGF
QVRSSTESPIDFVVDITISGQNTDEAQITQSTISRACKIVCDRNPYTAI.PAAGF
DSSEKNT.FLGEKRAKWKPNRDMGL.TYNGALVIMPRGSPFESQPGWRI.SVCGDY
TLRETSQAQGRKLVESQVNDLQDGLIDLCATLLWRDGL.FHTPTQKIBALQD
INARPCQVGLNTLAFPSINRKEVVEKQPMAYLSCGHYAGHNMCHSDTEANERE
CPMCRTPGPVPLMIGCEAGFPYDAGPRTAFTPCGHVSEKSAKXWSQI.PLPHGTHA
FHAACPATQLVGEONCTILITQGPID"

ORIGIN

Alignment Scores:
Pred. No.: 1,46e-166 Length: 1565
Score: 2290.00 Matches: 420
Percent Similarity: 100.0% Conservative: 0
Beet Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-041-030-4 (1-420) x AX817203 (1-1565)

QY 1 MeepheSerProgljngljugluhneCyalaProaenlysgluprovalyetyrly 20
Db 51 ATGTTTCCCTGGCCAGAGAACTGCGCCCAATAAGAGCCAGTAATACGGG 110
QY 21 GluleuValValleuglytyrfaenglyValaleuProaenglyabpaqglyaagaylys 40
Db 111 GAGCTGTGTGTCTCGGGTCAATAGTGTCTTCAATGAGATGAAAGACGAGGAAA 170
QY 41 SerarPhealaLeuTyrylYaragProlyealaAaenglyVallyrProserThValHis 60
Db 171 AGTGAATTTGCTCTCAAGAGGSCCAAGGCAATGCTGTCAACCAAGCACCTTCAT 230
QY 61 ValiIleSerThProglinaIleSerLyalaIleSerCyalygljnglInhiIleSer 80
Db 231 GTGATATCCAGCCCAAGGATCCAGGCTATGCTGCAAGAGGTCAACAGATATACC 290
QY 81 TyrThriLeuSeraragaenglnthValValValgluTyrrhIsaaplybaepThAsp 100
Db 291 TACACTTGTTCAGAGAAATCAGACTGTGTGTGAGTACACACATGATTAAGATACGAT 350
QY 101 MetPheGlnValGlyYarSerThgluSerProIleAspPheValValThraPThriLe 120
Db 351 ATGTTTAGGTGGCAATCAACAGAAACCTTATCCATCTTGTTCACAGACACATTT 410
QY 121 SerGlySerGlnAaenThraPgluaIleGlnIleThrgInserThriIleSerarPheAla 140
Db 411 TCTGGCAGCCAGAACACAGAGAACCCAGATCAACAGAGCACCATATCCAGGTTCCGC 470
QY 141 CyabArgIleValCyabapagaenglnProTyrrhIalargIlePheIalaleglyPhe 160
Db 471 TGCAGGTCTGTGTGCGACAGAAAGAACTTACACAGACGAGATATTGCGCGCGGATTT 530
QY 161 AspSerSerLyAaenIlePheLeuglygluValIalalValystrPlyAaenProaengly 180
Db 531 GACTCTTCAGAAACATATTTCTTGAAGAAAGCAGAGAAAGTGAAGAAACCCGACGCGC 590
QY 181 HisMetAspGlyLeuThraThraenglyValIleValMetHisPArOArglyglyPheThr 200
Db 591 CACATGATGGGCTCACTACTAATAGGCTCTGTGATGATCAACAGAGGGGCTTACCC 650
QY 201 GluGluSerGlnProgljvalTyraGluIleSerValCyaglyAaPvalTyrrhLeu 220
Db 651 GAGGAGTCCCGAGCCCGGGGTCTGCGCGAGATCTCTGTGTGAGATGTGTACACCTTG 710
QY 221 ArgGluThraArgSeraralaglnIalarglylyrleuValGluSerGluThraAaValleu 240
Db 711 CGAGAAACCGAGTGGCGCCAGCAAGAGAAAGCTGGGAGAAAGTGAAGCAACGCTCTG 770
QY 241 GlnAspGlySerLeuIleAspLeuCyaglyValaThriLeuLeuThraPThraIAspGly 260

Db 771 CAGAGCGGCTCCCTCATGACTGTGTGGGGCACTCTCTCTGAGAAACAGAGATGG 830
QY 261 LeuPheIsthrProThrglnlyshIsIleGluAlaleuArglnIleAaenAla 280
Db 831 CTTTTCATATCTCCAACTCAGAAACATTAAGAGCCCTCCGAGAGATTAACGCCGC 890
QY 281 ArgProGlnCyabProValGlyLeuAaenThriLeuAlaPheProSerIleAaenArglyglu 300
Db 891 CGGCTCAAGTCTCTGGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 950
QY 301 ValValGluGluIlysglnProThraIaTyrrLeuSerCyaglyIshValIsIlyTyrrHis 320
Db 951 GTGGTGAAGAGAGAGAGAGCCCTGGGCAATCTAGTGTGGCCACGTCGACGGGTCAC 1010
QY 321 AaThrGlyHisArgSerAaPThrgluAlaAaenglyluCyabPromeCyabArgThr 340
Db 1011 AACTGGGGCCATGGAGTACACGAGGCCAACAGAGGAGAGTCTCCATGTGCAGACT 1070
QY 341 ValGlyProTyrrValProLeuThriPleuglyCyaglyuAlaGlyPheTyrrValaPaIaGly 360
Db 1071 GTGGGGCCCTATGCTGCTCTGCTGGCTGTGAGAGAGATTTATGTAGACGAGGA 1130
QY 361 ProProThriAlaPheThraProCyaglyHisValCyabSerGluYsSerAlaIlyTyrr 380
Db 1131 CCGCCAACTCATGCTTCACTCCCTGTGACAGGTGTCTGGAGAAAGTCTGCAAAATAC 1190
QY 381 TrpSerGlnIleProLeuProIshIsglyThriAlaPheHisAlaIaCyabProPheCys 400
Db 1191 TGGCTCAGATCCCGTGTCTCAAGAACTCAAGCATTTACGCTGCTTCTGTGT 1250
QY 401 AlaThrglnLeuValGlygluGlnAaenCyalyLeuLeuIlePheGlnIlyProIleAsp 420
Db 1251 GCTACACAGCTGTGTGGGAGCAAACTGCATCAATTAATTTTCCAGGTCATTCATGAC 1310

RESULT 2
BC009476 1811 bp mRNA linear PRI 08-MAR-2005
LOCUS Homo sapiens pellino homolog 2 (Drosophila), mRNA (cDNA clone
DEFINITION MGC:15066 IMAGE:3942712), complete cds.
ACCESSION BC009476 GI:33872093
VERSION BC009476.2
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1811)
Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg
B, Buettow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
Max SI, Wang J, Hsieh P, Diatchenko L, Marusina K, Farmer AA, Rubin
GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL,
Scheetz TE, Brownstein WJ, Ustin TB, Toshlyuk S, Carninci P,
Prange C, Raha SS, Loggellano NA, Peters GJ, Abramson RD, Mullany
SJ, Bosak SA, McKernan PJ, McKernan KJ, Malek JA, Gunaratne PH,
Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW,
Villalón DK, Wuzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J,
Helton E, Kettelman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakeley RW,
Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,
Schmutz J, Myers RM, Butlerfield YS, Krzywicki MI, Skalska U,
Smalins DE, Schnerch A, Schein JE, Jones SJ and Maitra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
JOURNALS
PUBMED
REFERENCE 2 (bases 1 to 1811)
DIRECTOR MGC Project.
AUTHORS
TITLE
JOURNAL
Submitted (25-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

LOCUS	BD190328	5579 bp	DNA	linear	PAT 17-JUL-2003
DEFINITION	Elk1 phosphorylation related gene.				
ACCESSION	BD190328				
VERSION	BD190328.1	GI:33000067			
KEYWORDS	WO 03008589-A/68.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 5579)				
AUTHORS	Masuzaki,O., Matsuda,A., Nagano,Y. and Suzuki,I.N.				
TITLE	Elk1 phosphorylation related gene				
JOURNAL	Patent: WO 03008589-A 68 30-JAN-2003;				
	ASHKI KASAI CORP.OSAMU MATSUZAKI,AKIO MATSUDA,YUKIKO NAGANO, NAOMI SUZUKI				
COMMENT	OS Homo sapiens (human)				
	PN WO 03008589-A/68				
	PD 30-JAN-2003				
	PF 15-JUL-2002 WO 2002JP007174				
	PR 18-JUL-2001 JP 01P 218204,31-AUG-2001 JP 01P 263450 PR				
	21-JAN-2002 JP 02P 012176				
	P1 OSAMU MATSUZAKI,AKIO MATSUDA,YUKIKO NAGANO,NAOMI SUZUKI PC				
	C12N15/54,C12N9/12,G01N33/50,G01N33/15,C07K16/40,A61K39/385, PC				
	A61K31/711				
FEATURES	CC Elk1 phosphorylation related gene				
	FH Key Location/Qualifiers				
	FT CDS Location/Qualifiers (177)..(1436).				
FEATURES	Location/Qualifiers				
source	1..5579				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Alignment Scores:					
Pred. No.:	6,286-166	Length:	5579		
Score:	2290.00	Matches:	420		
Percent Similarity:	100.0%	Conservative:	0		
Best Local Similarity:	100.0%	Mismatches:	0		
Query Match:	100.0%	Indels:	0		
DB:	6	Gaps:	0		
US-10-041-030-4 (1-420) x BD190328 (1-5579)					
QY	1	MetPheSerProGluGlnGluGlnHisGlyAlaProAsnLysGluProValLysTrpGly	20		
DB	177	ATGTTTTCCTCCCTGGCCAGAGGAAACATGCGCCCAATTAAGAGCCAGTGAATAACGGG	236		
QY	21	GluLeuValValLeuGlyTrpAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgLys	40		
DB	237	GAGCTGGTGGTGGCTGGGTTACATGCTCTTACCCCAATGAGATAGAGACGACGAGAA	296		
QY	41	SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis	60		
DB	297	AGTAGATTTCCTCTTACAGCGGCCCAAGGCAATGTGTCAACACCAGACCCGTCCAT	356		
QY	61	ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlnGlnHisSerLysSer	80		
DB	357	GTCATATCAGCCGCCAGGACATCCAAAGCTATCAGCTGCAAAAGTCAACACAGTATATCC	416		
QY	81	TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAspLysAspThrArgP	100		
DB	417	TACACTTTGTCAAGGAATCAAGACTGTGGTGGTGGAGTACACACATGATTAAGATTCGGAT	476		
QY	101	MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValValThrAspThrIle	120		
DB	477	ATGTTTCAGGTGGCGCAGATCAACAAAGACCTATTCGACTTCGTTGTCAACAGACGATT	536		
QY	121	SerGlySerGlnAspThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla	140		
DB	537	TCTGGCACCCAGAAACGACGACCAAGCCCAAGATCACACAGACCAATATCCAGGTTCCGC	596		

QY	141	CysArgIleValCysAspArgAsnGluProCysThrAlaArgIlePheAlaIleGlyPhe	160
Db	597	TGCAGATCGTGTGCAGACGAGAAAGAACTTACACAGACGAGATATTCGCCCGGAAATT	656
QY	161	AspSerSerIysAsnIlePheLeuGlyGluIleValAlaIleAspTrpLysAsnProAspGly	180
Db	657	GACCTTCCAAAACATATTTCTTGAGAAAAGGACCAAAAGTGGAAAACCCGACGGC	716
QY	181	HisMetAspGlyLeuThrTrpAsnGlyValIleuValMetHisProArgGlyGlyPheThr	200
Db	717	CACATGATGGGCTCACACTACTAAATGGGGCTCTGGTGTATGCATCCACAGAGGGGGCTTAC	776
QY	201	GluGluSerGlnProGlyValAlaTrpArgGluIleSerValCysGlyAspValTyrThrLeu	220
Db	777	GAGAGTCCACGCCCGGGGCTGGCGGAGATCTCTGTCTGTGGAGATGGTACCTTG	836
QY	221	ArgGluThrArgSerAlaGlnGlnAlaArgGlyLysLeuValGluSerGluThrAsnValIleu	240
Db	837	CGAGAAACAGAGTGGGCCAGCAACAGAAAGCTGGTGGAAATGAGCCAACTGCTCTG	896
QY	241	GlnAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuLeuTrpArgThrAlaAspGly	260
Db	897	CAGACGGCTCTCCATTCATTCCTGTGGGGCCACTCTCTCTGGAGAAACAGCAGATGGG	956
QY	261	LeuPheHisThrProThrGlnLysHisIleGluIleLeuArgGlnGluIleAsnAlaAla	280
Db	957	CTTTTCACTATCCAACTCAGAACCACTATGAAACCTCCGGCAGAGATTTAACGGCCG	1016
QY	281	ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu	300
Db	1017	CGGCTCAGTGTCTCTGTGGGGCTCAACACCCTGGCTTCCCAAGCATCAACAGAAAGAG	1076
QY	301	ValValGluGluLysGlnProTrpAlaTyrIleuSerCysGlyHisValHisIleGlyTyrHis	320
Db	1077	GTGGTGAGAGAGAGACAGCCCTGGGCAATCTCAGTTGGGCCACGTGGACAGGGTATCAC	1136
QY	321	AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr	340
Db	1137	AACGGGGGCCATCGGAATGACACAGAGGCCAACAGAGGGAGTGTCCATGTGCAAGACT	1196
QY	341	ValGlyProTyrValProLeuTrpLeuGlyCysGlyAlaGlyPheTyrValAspAlaGly	360
Db	1197	GTGGGCCCTTATGGCTCTCTGGCTTGGCTGTAGGACGATTTTATGTAAGCGCAGGA	1256
QY	361	ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr	380
Db	1257	CCGCCAATCAATGCTTCACTCCCTGTGGACACGTGTCTGGAGAAATGTGCAAAATAC	1316
QY	381	TrpSerGlnIleProLeuProHisIleGlyThrHisAlaPheHisAlaAlaCysProPheCys	400
Db	1317	TGGTCTCGAATCCCGTGTGCTTCATGAAATCATATTCATTCACGCTGTCTGCCCTTCTGT	1376
QY	401	AlaThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp	420
Db	1377	GCTACACAGCTGTGGGGAGCAAAATCTGCATCAATTAATTTTCCAAAGTCCAAATTGAC	1436
RESULT 4			
AF302502	5597 bp	mRNA	linear
LOCUS			PRI 05-JUL-2001
DEFINITION	Homo sapiens pellino 2 (PELL2) mRNA, complete cde.		
ACCESSION	AF302502		
VERSION	AF302502.1	GI:10242352	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
TITLE	Homnidae; Homo.		
	1 (bases 1 to 5597)		
	Reisch, K., Jockusch, H. and Schmitz-John, T.		
	Assignment of homologous genes, Pell1/PBL1 and Pell2/PBL2, for		
	the pelle adaptor protein Pellino to mouse chromosomes 11 and 14		
	and human chromosomes 2p13.3 and 14q21, respectively, by physical		

JOURNAL
PUBMED
REFERENCE
AUTHORS
JOURNAL
TITLE

and radiation hybrid mapping
Cytogenet. Cell Genet. 92 (1-2): 172-174 (2001)
11306823
2 (bases 1 to 5597)
Resch, K., Jockusch, H. and Schmitt-John, T.
Direct Submision
Submitted (05-SEP-2000) Developmental Biology and Molecular
Pathology, University of Bielefeld, Universitaetsstrasse 25,
Bielefeld, NRW 33615, Germany

FEATURES

Source

Location/Qualifiers
1..5597

/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"

1..5597
/gene="PEL12"

177..1439
/gene="PEL12"

/note="similar to Drosophila melanogaster Pellino"

/product="pellino 2"

/protein_id="PAG15390.1"

/db_xref="GI:10242353"

/translation="MPSGQSEHCAPNKEPVKGYELVVLGNGALPMDGRKRKSRPA
LYRPRKANGVKSPTVHYITSPQASKISCKGSHISYTLRNQVVEYTHDDTDMF
QNGRSTSPIDPVVTDTISSQNTDEAQITOSTISPRACIVCDRNEPTARIIPAGP
DSSNLTLEKAKAKNPDGMDGLTNGVLMHPRGFTBESQGVWRIKISVCGDY
TLRTSRAGQKGLVESITVLDGSLIDCGATLMTADGLFHTPTQKRIEALROE
INARPCQPVGLMTLAPPSINRKEVVEKQPMWLTSCGHGYHNMHREDTEINERE
CPMCRITGVPLVPLMGAEAGPYVDAGPPTAFPCGHVCSSEKAKYWSQIPLEGTHA
FNAACPCPAQVLQVEQNCIKLIPGSPID"

ORIGIN

Alignment Scores:

Pred. No.: 6.3e-166 Length: 5597
Score: 2290.00 Matches: 420
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-041-030-4 (1-420) x AF302502 (1-5597)

QY 1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
DB 177 ATGTTTCCCTGGCCGAGAGAACACTGGCCCCCAATAGAGCCAGCAATATACGGG 236
QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGly 40
DB 237 GAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 296
QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
DB 297 AGTAGATTTGCTCCCTCAAGCGGCCCAAGGCAATGGTGTCAACCAAGACCGTCCAT 356
QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
DB 357 GTGATATCCACGCCCCAGGATCCAGAGCTATCAGTCGCAAGAGGTCAACACATATATCC 416
QY 81 TyrThrIleSerArgAsnGlnThrValValValGluTyrThrHisAspLysAspThrAsp 100
DB 417 TACCTTTGTCAGAGAAATCAGACTGTGTGTGGTAGTACACACATVTAAGATACGGAT 476
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
DB 477 ATGTTTCAAGGTGGCCAGATCAACGAAAGCCCTATCCACTTCGTGTCAAGACAGAT 536
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 537 TCTGGCCGACCAAGACGAGAGCCAGATCAACAGAGACACATATCCAGGTGGCC 596
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrThrAlaGlyIlePheAlaIleGlyPhe 160

DB 597 TGACAGATCGTGTGCGACAGAAATGAACCTTACACAGACGAGATATTCGCCGCCGATTT 656
QY 161 AspSerLysAsnIlePheLeuGlyGlnLysAlaAlaLysTyrLysAsnProAspGly 180
DB 657 GACTCTTCCAAAACATATTTCTTGGAGAAAGCCAGCAAGAGGAAAAACCCCGACGGC 716
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB 717 CACTGATGGGCTCACTACTAATGCGCTCTGTGTATGATCATCAGAGGGGCTTACC 776
QY 201 GluGluSerGlnProGlyValITPArgGluIleSerValCysGlyAspValTyrThrLeu 220
DB 777 GAGAGTCCAGACCCGGGCTCGGCGGAGATCTGTGTGTGAGATGTATACCTTG 836
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
DB 837 CGAGAAACAGGTCTGGCCGACGACAGAGAAAGCTGTGAGAAAGTGAACCAACGTCCTG 896
QY 241 GluAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuLeuTPArgThrAlaAspGly 260
DB 897 CAGAGCGGCTCCCTCATTAACCTGTGTGGGCGCACTCTCTGAGAGACAGCATGGG 956
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAla 280
DB 957 CTTTTCATATCTCACTCACTCAAGAGCATAGAGCCCTCGGACAGAGATTAACGCCGCC 1016
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
DB 1017 CGGCTCATGTCTCTGTGGGCTCAACACCTGGCTTCCAGCATCAACAGAAAGAG 1076
QY 301 ValValGluGluLysGlnProTPAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
DB 1077 GTGGTGGAG 1136
QY 321 AsnTPGlyHisArgSerAspThrGlnAlaAsnGluLysGluCysProMetCysArgThr 340
DB 1137 AACCTGGGCGCATGGAGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGACT 1196
QY 341 ValGlyProTyrValProLeuTPLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
DB 1197 GTGGGCCCTCATGTCTCTGTGGGCTCAACACCTGGCTTCCAGCATCAACAGAAAGAG 1256
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
DB 1257 CCGCAACTATGCTTTCATCTCTTGAACAGGTGTGAGAGAGAGAGAGAGAGAGAGAG 1316
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
DB 1317 TGGTTCAGATCCCGTGGCTCATGAACTCATGATTTTCAACGCTGCTTCTGTGT 1376
QY 401 AlaThrGlnLeuValGlyGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
DB 1377 GCTACACAGCTGGTGGGAGAGAAACCTGCATCAATATTTTCCAGAGTCCAAATTGAC 1436

RESULT 5
AX686405 5921 bp DNA linear PAT 29-MAR-2003
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE
1 Powers, S., Mu, D., Xiang, P. and Peng, Y.
Diagnosis and treatment of cancer using mammalian pellino
polypeptides and polynucleotides
Patent: WO 02059611-A 3 01-AUG-2002;
JOURNAL
TITLRS
AUTHORS
Polarik Inc. (US)
Location/Qualifiers

source
1. 5921
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
501..1763
/note="unnamed protein product; human pellino 2"
/protein_id="CAD86576.1"
/db_xref="GI:29372144"
/translation="MPSPGQEHCPAPKEPVKYGELVVLGNGALPNDGRGRKRPAL
LYRPRKNGVKSPTVHYISTPQSKASISCGHSISTLSRNTQVVEVTHDQTMF
QVNRSTSPIDFVVTDTISGQNDQITSTISPRACKICDREPEYRARIAPAGF
DSSKNIPLGKAAKKNPDGMDGLTNGLVMMPRGFTPEQPGVREIISVCGDY
TLRETSQOQRGLVESSEVNLQDGLIDLCGALTMLRTLDGLFHTPTQGLALOE
INARPCPGVGLTAPPSINRKEVEKQPMAYLSCGHYGLNMGHRSIDTANERE
CPMCRITGPPVPLMIGCEAGFYVDAGBPRTAFTPCGHVSEKSAKXWSQILPHGHTA
FHACPCPATQIYVEQNCITLITQSPID"

ORIGIN

Alignment Scores:

Pred. No.:	6,72e-166	Length:	5921
Score:	2230.00	Matches:	420
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-041-030-4 (1-420) x AK686405 (1-5921)

QY 1 MetPheSerProGlyngInGluGlnHsCyAlaProAbnlysgInuProValyTyrgly 20
Db 501 ATGTTTCCCCCTGGCCAGAGAAACCTGGCCCCCAATAGAGCCAGGAAATACGGG 560
QY 21 GluLeuValValleuGlyTyraenGlyAlaLeuProAbnlygaPaArgGlyATGARGly 40
Db 561 GAGCTGGTGGTGGCTGGGATCAATGTGCTTACCCATGAGATGAGAGAGAGAGAAA 620
QY 41 SerArgPheAlaLeuTyrylYsArgProLyAlaAsnGlyVallyProSerThValHis 60
Db 621 AGTAGATTTCCTCTCAAGAGCGGCCCAAGGCAATGTGTCAAAACCCAGCACCGTCCAT 680
QY 61 ValIleSerThProGlnAlaSerLyAlaIleSerCyAllysglyGlnHsSerIleSer 80
Db 681 GTGATATCCAGCCGCCAGGATCCAGAGCTATCAGTCAAGAGGTCAACACAGATATCC 740
QY 81 TyThrIleuSerArgaenGlnThValValGlnTyThrHisaaplysaapThraSp 100
Db 741 TACACTTTCAGAGAAATCAGACTGTGTGGTGAATACACACTGATTAAGATACGAT 800
QY 101 MetPheGlnValGlyATGSerThGlySerProIleAspPheValThraapThrIle 120
Db 801 ATGTTTCAGGTGGGCAATCAACAGAAAGCCCTTACACTTCCTGTGCACAGACACAT 860
QY 121 SerGlySerGlnaenThraapGlyAlaGlnIleThrgInserThrIleSerArgPheAla 140
Db 861 TCTGGCAGCCAGAAACGAGAGCCAGATCAACAGAGACACCATATCCAGGTGGCC 920
QY 141 CyAlaArgIleValCyAsaPaArgaenGlnuProTyThrAlaArgIlePheAlaIleGlyPhe 160
Db 921 TCGAGAGTCTGTGCGACAGAAATGAACCTTACACACAGAGATATTCGCCCGGATTT 980
QY 161 AspSerSerLyAsnIlePheLeuGlyGlyAlaIleValyThraPlyAsnProAspGly 180
Db 981 GACTCTTCCAAAACATATTTCTTGGAGAAAAGCAGCAAAAGTGAAGAAACCCGACGCG 1040
QY 181 HisMetAspGlyLeuThraenGlyValleuValMetHisProArgGlyGlyPheThr 200
Db 1041 CACATGATGGGCTCATCTAATGCGCTCTGTGTATGATCATCACAGAGGGGCTTACC 1100
QY 201 GluGlnSerGlnProGlyValThraArgGlnIleSerValCyAsGlyAspValTyThrLeu 220
Db 1101 GAGGAGTCCAGCCGGGGTCTGGCGGAGATCTCTGTGTGAGATGTGTACACCTTG 1160

QY 221 ArgGluThraArgSerAlaGlnGlnArgGlyTyraenValGluSerGlnThraenValleu 240
Db 1161 CGAGAAACAGAGTCCGCGCCAGAACGAGAAAGCTGGTGAAGTGAACCAACGTCTTG 1220
QY 241 GluAspGlySerIleuIleAspLeuCyGlyAlaThrLeuLeuThraArgThraIleAspGly 260
Db 1221 CAGAGCGCTCCCTCATTAACCTGTGTGGGCGACCTCTCTGAGAGAACAGAGATGGG 1280
QY 261 LeuPheIleThraProThrgInlyShsIleGlyAlaLeuArgGlnIleuIleAsnAla 280
Db 1281 CTTTTCATCTCTCAACTCAAGAGCATTAAGAACCTCCGAGAGATTAACGCCGCC 1340
QY 281 ArgProGlnCyAsProValGlyLeuAenThraLeuAlaPheProSerIleAsnArglyGly 300
Db 1341 CGGCTCAGTGTCTGTGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1400
QY 301 ValValGlnGlyLeuProTyThrAlaTyrylLeuSerCyGlyHisValHisGlyTyHis 320
Db 1401 GTGGTGAAGAGAGAGAGCCCTGGGCAATCTCAGTGTGGCCACGTCACGGGTACAC 1460
QY 321 AsnTrpGlyHisArgSerAspThrglyAlaAsnGlyArgGlyCysPProMetCyAsArgThr 340
Db 1461 AACTGGGCGCATCGAGTGAACAGAGGCCAAGAGAGAGTGTCCATGTGCAGAGACT 1520
QY 341 ValGlyProTyrylValProLeuThraLeuGlyCyGlyAlaGlyPheTyrylAlaPaArgly 360
Db 1521 GTGGGCGCCATGTGCTCTCTGGCTGTGGTGAAGAGATTTTATGTAGAGCGAGGA 1580
QY 361 ProProThraHisAlaPheThraProCyGlyHisValCySerGlyLySerAlaTyryl 380
Db 1581 CCGCAACTCATCTCTTCACTCCCTGTGACAGGTGTGCTCGAGAGAGTCTGCAAAATAC 1640
QY 381 TrpSerGlnIleProLeuProHisGlyThraHisAlaPheHisAlaIleCysProPheCys 400
Db 1641 TGCTTCAGATCCCGTTCCTCATGAACTCATGCACTTTCACCTGCTTCCCTTCTGT 1700
QY 401 AlaThrgInleuValGlyGlnGlnAenCyAlaLeuLeuIlePheGlnGlyProIleAsp 420
Db 1701 GCTACACAGCTGTTGGGAGAGAAAGCTGATCAATTAATTTTCCAGGTCCAAATTGAC 1760

RESULT 6
AR483568 1263 bp DNA linear PAT 14-MAY-2004
LOCUS
DEFINITION
ACCESSION AR483568
VERSION AR483568.1 GI:47246652
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1263)
AUTHORS
TITLE
HUMAN PELLINO POLYPEPTIDES
JOURNAL
Patent: US 6703487-A 7 09-MAR-2004;
Immunex Corporation; Seattle, WA;
MOX;

FEATURES
source location/Qualifiers
1..1263
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:	3.94e-166	Length:	1263
Score:	2283.00 <td>Matches:</td> <td>419</td>	Matches:	419
Percent Similarity:	99.8% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	99.8% <td>Mismatches:</td> <td>1</td>	Mismatches:	1
Query Match:	99.7% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	6	Gaps:	0

US-10-041-030-4 (1-420) x AR483568 (1-1263)

QY 1 MetPheSerProGlyngInGluGlnHsCyAlaProAbnlysgInuProValyTyrgly 20


```

Db      1 ATGTTTCCCTGGCCAGAGAAACACTGGCCCCCAATAGAGAGCCAGTGAATACGGG 60
Qy      21 GtluLeuValValleuGlyYrAenglyValaleuProaenglyAaparglyAargly 40
Db      61 GACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Qy      41 SerArgPheAlaLeuYrlyAargProlyAalaaenglyVallyAapProserThrValHis 60
Db      121 AGTAGATTCCCTCTACAGAGCCGCCAGAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Qy      61 ValIleSerThrProGlnAlaSerlyAalIleSerCyslySgIyGlnHisSerIleSer 80
Db      181 GTGATATCCAGCCGCCAGGATCCAGAGCTATGACGTGCAAGAGTCAACAGATATCC 240
Qy      81 TyrThrLeuSerArgaengIntThrValValleuGlyYrThriAaplyAapThraAp 100
Db      241 TACACTTTGTCAAGAGATCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy      101 MetPheGlnValGlyAargSerThrGlnUserProIleAapPheValValThraApThrIle 120
Db      301 ATGTTTCAGGTGGGACATCAACAGAAAGCCCTATGCACTTCTGTGCACAGACAGATT 360
Qy      121 SerGlySerGlnAanThraAapGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361 TCTGGCAGCCAGAAACAGAGCAAGCCAGATCAACAGAGACCATATCCAGGTTCC 420
Qy      141 CysArgIleValCysAapArgaengIntProYrThraIlaargIlePheAlaIleGlyPhe 160
Db      421 TGCAGAGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Qy      161 AapSerSerlyAanIlePheLeuGlyVallyAalIleAlySTPlyAanProAapGly 180
Db      481 GACTCTTCCAAAACATATTTCTTGTGAGTAAAGGACAAAGTGGAACCCCGACGCGC 540
Qy      181 HiMetAapGlyLeuThrThraaenglyValleuValMetHisPProAarglyGlyPheThr 200
Db      541 CAATGATGGGCTCACTACTATATGCGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy      201 GlnGlnUserGlnProGlyValITPArgGlnIleSerValCysGlyAapValYrThrLeu 220
Db      601 GAGGAGTCCAGCCCGGGGTCTGGCCGAGATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Qy      221 ArgGlnThraArgSerAlaGlnGlnAarglyValleuValGlnUserGlnThraaenValleu 240
Db      661 CGAGAAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Qy      241 GlnAapGlySerLeuIleAapLeuCysGlyValaThrLeuLeuThraArgThraIlaapGly 260
Db      721 CAGGACGGCTCCCTCATTTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy      261 LeuPheHisThraProThrglnlyAalIleGlnValaLeuAargGlnIleAaenAla 280
Db      781 CTTTTCATATCTCAACTCAGAGACCATAGAACCCCTCCGGCAGAGATTAAAGCGCC 840
Qy      281 ArgProGlnCysProValGlyLeuAanThrLeuAlaPheProserIleAaenAlySgln 300
Db      841 CGGCTCAGAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Qy      301 ValValGlnGlnlySglnProThraIaYrleuSerCysGlyHisValHisIleGlyYrHis 320
Db      901 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Qy      321 AaenITrGlyHisArgSerAapThrglnAlaAaenglyValGlnCysPrometCysArgThr 340
Db      961 AACTGGGGCCATGTGAGTGAACAGAGGCCAAGAGAGAGTGTCCCATGTGAGACT 1020
Qy      341 ValGlyProYrValProLeuITPLeuGlyCysGlnAlaGlyPheYrValAaapAlaGly 360
Db      1021 GTGGGCCCCATATGTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Qy      361 ProProThraIlaPheThraProCysGlyHisValCysSerGlnlySserAlaIleYr 380

```

```

Db      1081 CCGCAACTCATGCTTTCACCTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Qy      381 TrpSerGlnIleProLeuProHisGlyYrThriAalPheHisAlaIleAlyCysProPheCys 400
Db      1141 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Qy      401 AlaThrGlnLeuValGlyGlnGlnAanCysIleLeuLeuIlePheGlnIlyProIleAap 420
Db      1201 GCTACACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260

RESULT 7
LOCUS      AX298079             1263 bp      DNA      linear      PAT 26-NOV-2001
DEFINITION Sequence 7 from Patent WO0183739.
ACCESSION  AX298079
VERSION     AX298079.1 GI:17128160
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1
            Bird, T.A. and Cosman, D.J.
            Human pellino polypeptides
            Patent: WO 0183739-A 7 08-NOV-2001;
            IMMUNEX CORPORATION (US)
FEATURES
            source          1..1263
                                /organism="Homo sapiens"
                                /mol_type="unassigned DNA"
                                /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      3,94e-166      Length:      1263
Score:          2283.00      Matches:      419
Percent Similarity: 99.8%      Conservative: 0
Best Local Similarity: 99.8%      Mismatches: 1
Query Match:    99.7%      Indels:      0
                        Gaps:      0
US-10-041-030-4 (1-420) x AX298079 (1-1263)
Qy      1 MetPheSerProGlyGlnGlnGlnHisCysAlaProAaenlySglnProVallySTYrGly 20
Db      1 ATGTTTCCCTGGCCAGAGAAACACTGGCCCCCAATAGAGAGCCAGTGAATACGGG 60
Qy      21 GtluLeuValValleuGlyYrAenglyValaleuProaenglyAaparglyAargly 40
Db      61 GACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Qy      41 SerArgPheAlaLeuYrlyAargProlyAalaaenglyVallyAapProserThrValHis 60
Db      121 AGTAGATTCCCTCTACAGAGCCGCCAGAGCAATGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Qy      61 ValIleSerThrProGlnAlaSerlyAalIleSerCyslySgIyGlnHisSerIleSer 80
Db      181 GTGATATCCAGCCGCCAGGATCCAGAGCTATGACGTGCAAGAGTCAACAGATATCC 240
Qy      81 TyrThrLeuSerArgaengIntThrValValleuGlyYrThriAaplyAapThraAp 100
Db      241 TACACTTTGTCAAGAGATCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy      101 MetPheGlnValGlyAargSerThrGlnUserProIleAapPheValValThraApThrIle 120
Db      301 ATGTTTCAGGTGGGACATCAACAGAAAGCCCTATGCACTTCTGTGCACAGACAGATT 360
Qy      121 SerGlySerGlnAanThraAapGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361 TCTGGCAGCCAGAAACAGAGCAAGCCAGATCAACAGAGACCATATCCAGGTTCC 420
Qy      141 CysArgIleValCysAapArgaengIntProYrThraIlaargIlePheAlaIleGlyPhe 160

```

Db 421 TGCAGATCGTGGCAGAGAACTTACACAGAGATTCGCGCGGATTT 480
Qy 161 AspSerSerlysaenllepheleuglyglulysalalalysrtrplysaenproaspoly 180
Db 481 GACTCTTCCAAAACATATTTCTTGAGATTAAGGACCAAAAGTGGAAAACCCCGAGGCC 540
Qy 181 HsMetAspGlyLeuThrTrpAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 541 CACATGATGGGCTCACTACTAATGGCGTCTGGTATGATGATCCAGAGGGGGCTTACC 600
Qy 201 GluGluSerGlnProGlyValTrpArgGluIleSerValCysGlyAspValTyrThrLeu 220
Db 601 GAGAGAGCCAGCCCGGGGTCTGGCGGAGATCTCTCTGTGGAGATGGTACACCTTG 660
Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnValLeu 240
Db 661 CGAGAAACAGGCTCGCCACAGCAACAGAAAGCTGGTGGAAAGTGAACCAACGCTCTG 720
Qy 241 GluAspGlySerLeuIleAspLeuCyseGlyValAlaThrLeuLeuTrpArgThrAlaAspGly 260
Db 721 CAGGAGCGCTCCCTCACTTGAACGTGTGGGGCCACTCTCTCTGGAGAACAGAGATGGG 780
Qy 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
Db 781 CTTTTCATCTCCACTCAGAACACATGAAAGCCCTCCGCGAGATTAACGCGCC 840
Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
Db 841 CGGCTCAGTGTCTGTGGGCTCAACACCCCTGCTCCAGCATCAACAGAAAGAG 900
Qy 301 ValValGluGluLysGlnProTrpAlaTyrLeuSerCysGlyHisValIleGlyTyrHis 320
Db 901 GTGGTGAAGAGAACGAGCCCTGGGCATATCTCATGTTGTGGCCACGTGCAGGGTACAC 960
Qy 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
Db 961 AACTGGGGCCATCGAGTGAACCGAGGCCAACAGAGGAGTGTCCCATGTGACGAGACT 1020
Qy 341 ValGlyProTrpValProLeuTrpLeuGlyCysGluAlaGlyPheTyrValAlaAspAlaGly 360
Db 1021 GTGGGCCCCATGCTCTCTGTGGCTGTGGAGGAGATTTATGTAGACGGCGGA 1080
Qy 361 ProProThrHisAlaPheThrProCysGlyHisValAlaCysSerGluLysSerAlaLysTyr 380
Db 1081 CCGGCAACTATGCTTTCATCTCCCTGTGACACGTGTCTGGAGAGTCTGCAAAATAC 1140
Qy 381 TrpSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1141 TGGTCTCAGATCCGTTGCTCATGAACTCATTCATTCACGCTGCTTCTCTGT 1200
Qy 401 AlaThrGlnLeuValGlyGluGlnAsnCysGlyLeuLeuLeuIlePheGlnGlyProIleAsp 420
Db 1201 GCTACACAGCTGTGGGAGCAAACTGCATCAATTAATTTCCAAAGTCCCAATTGAC 1260

RESULT 8
AR483567 1260 bp DNA linear PAT 14-MAY-2004
LOCUS AR483567
DEFINITION Sequence 5 from patent US 6703487.
ACCESSION AR483567
VERSION AR483567.1 GI:47246651
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1260)
AUTHORS Bird, T.A. and Cosman, D.J.
TITLE Humanellino polypeptides
JOURNAL Patent: US 6703487-A 5 09-MAR-2004;
Immunex Corporation; Seattle, WA,
MOX;
FEATURES
source Location/Qualifiers
1..1260

ORIGIN /organism="unknown"
/mol_type="genomic DNA"
Alignment Scores:
Pred. No.: 7,216-158 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.4% Conservative: 10
Best Local Similarity: 95.0% Mismatches: 10
Query Match: 95.0% Indels: 1
DB: 6 Gaps: 1
US-10-041-030-4 (1-420) x AR483567 (1-1260)
Qy 1 MetPheSerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
Db 1 ATGTTTTCCTCCGCGCCAGAGAGAACCCACCGCCCAACAGAGAGCGGTGAATTCGGG 60
Qy 21 GluLeuValValLeuGlyTyrAsnGlyValAlaLeuProAsnGlyValAspArgGlyArgArgLys 40
Db 61 GAGCTGTGGTCTCTGGGTATCAATGTGTCTTACCTAATGTGTACAGGGGCGAGAGAA 120
Qy 41 SerArgPheAlaLeuTyrTrpAsnArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 121 AGCAGATTGGCTCTTAAAGCGGACCTACGCGCAAGTGTGTCAAAACCCAGCAAAATCCAC 180
Qy 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
Db 181 ATGCTTCCACACCAACAGCGCTCCAGGCCATCTACAGAGACATCACAGCATATTCG 240
Qy 81 TyrThrLeuSerArgAsnGlnThrValValValGluTyrThrHisAspLysAspThrAsp 100
Db 241 TACACGTTGTCAAGGAGCCAGACGATAGTGTGATGATACACACATTAAGACACGAGAC 300
Qy 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
Db 301 ATGTTTCAGTGGGAGGTCAACAGAAAGCCCATTTGACTCTGTGTCAACAGACGGGT 360
Qy 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCCGGCGCTCAACCAAGAT---GCCAGATCACACAGACCACTCTCTAGTTCGCA 417
Qy 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 418 TGCAGATCGTGTGTACAGAAACGACCATATACAGACGATATTCGCGGAGGATTC 477
Qy 161 AspSerSerlysaenllepheleuglyglulysalalalysrtrplysaenproaspoly 180
Db 478 GATTTCTCCAAAATATCTTCTTGTGAGAGAAAGCAGAAATGGAAAACCTGTATGGA 537
Qy 181 HsMetAspGlyLeuThrTrpAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 538 CACATGATGAGACTCACTACCAATGATGCTCTAGTATGACCCCGAGAGAGGCTTACC 557
Qy 201 GluGluSerGlnProGlyValTrpArgGluIleSerValCysGlyAspValTyrThrLeu 220
Db 558 GAGGAATCCAGGCTGAGAGTCTGAGAGAGATCTCTGTGTGGATGTATCACTTG 657
Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnValLeu 240
Db 658 CGAGAGACCAAGTCCGCGCCAGAGAGGAGAAAGCTGTGGAAAGTGAAGACCAACGCTCTG 717
Qy 241 GluAspGlySerLeuIleAspLeuCyseGlyValAlaThrLeuLeuTrpArgThrAlaAspGly 260
Db 718 CAGAGCGCTCCCTCACTTGAACGTGTGTGGGGCCACTCTCTCTGGAGAACCGCAGATGGC 777
Qy 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaAla 280
Db 778 CTTTTCAGGCTCTACTCAGAAAGCATAGAAAGCCCTCCGAGAGAGATCAATGACAGCC 837
Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
Db 838 CGACCCAGTGCCTGTGGGCTTAAACACCTGTGCTTCCCGAGCATCAACCGAAAGGAA 897

QY	301	ValValGluGluValValGluProTProTalaTyrlLeuSerCySGlValHISglTyrlHis	320
Db	898	GTGTGGAGAGAAAGACACCCCTGGGCAATACCTGAGCTCGGCATGTGCACGGCTCAAC	957
QY	321	AenTPrpGlyHISArgSerASPThrGluValaangluValArgGluCySPrometCysArgThr	340
Db	958	AGCTGGGGCCATCGGAGCGACCGCGGAACCCAAACAGAGAGGTGTCCATGTGCAGACT	1017
QY	341	ValGlyProTyrlValProLeuTPrLeuGlyCySGluValaglyPheTyrlValAspAlagly	360
Db	1018	GTGGGCCCCCTACGTCCTCTCTGTGGCTGTGTGAGGAGATTTTAATGTGATCGCGGA	1077
QY	361	ProProThrHISalaPheThrProCySGlyHISValCySserGluValSerAlaValTyrl	380
Db	1078	CCCCCACTCACGCTTTACACCCCTCGCGGACAGCTGTTCAGANAAGTCTGCCAAGTAC	1133
QY	381	TrpSerGlnIleProLeuProHISglTyrlHISalaPheHISalaAlaCySProPheCyS	400
Db	1138	TGTGTGCGAGATCCCACTGCCACCGAGGAACGACGGCTTTCATGCGCGCTGTCTGTGC	1197
QY	401	AlaThrGlnLeuValGlyGluGlnAsnCySlelyLeuIlePheGlnGlyProIleAsp	420
Db	1198	GCCACGACGCTGTGTGTGGAAACAGAACATCAATGATTTTCCAAAGTCCAGTGCAC	1257

RESULT 9

AX298077

LOCUS AX298077 1260 bp DNA linear PAT 26-NOV-2001

DEFINITION Sequence 5 from Patent WO0183739.

ACCESSION AX298077

VERSION AX298077.1 GI:17128159

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1. Bird, T.A. and Cosman, D.J. Human pellino polypeptides Patent: WO 0183739-A 5 08-NOV-2001; IMMUNEX CORPORATION (US)

FEATURES

source location/Qualifiers

1..1260 /organism="Mus musculus" /mol_type="unassigned DNA" /db_xref="taxon:10090"

ORIGIN

Alignment Scores:

Pred. NO.:	7,21e-158	Length:	1260
Score:	2175.50	Matches:	399
Percent Similarity:	97.4%	Conservative:	10
Best Local Similarity:	95.0%	Mismatches:	10
Query Match:	95.0%	Indels:	1
DB:	6	Gaps:	1

US-10-041-030-4 (1-420) x AX298077 (1-1260)

QY	1	MetPheSerProGlyGlnGluGlnHISCySalaProAenlyGluProValIleTyrlGly	20
Db	1	ATGTTTTCCTCCCGGCGCAGAGAGAACCCAGCGCCCCCAACAGAGACCGGTGAATAACGGG	60
QY	21	GluLeuValIleValLeuGlyTyrlAsnGlyAlaLeuProAenGlyAspArgGlyValArgIleGly	40
Db	61	GAGCTGTGTGTCTCGGGTACATGTGTCTTACTTAATGTGAACAGGGGCGAGGAGAA	120
QY	41	SerArgPheAlaLeuTyrlIleArgProIleAlaIleGlyValIleProSerThrValHIS	60
Db	121	AGCAGATTTCCTCTATTAAGCGGACCTACGCGCAGTGTGTCTCAACCCAGACCAATTCAC	180
QY	61	ValIleSerThrProGlnAlaSerIleValIleSerCySlySGlyGlnHISerIleSer	80
		

Db	181	ATGATCTCCACACACAGAGGCGTCCAAAGGCCATCAAGCTCCAGAGGACATCAACGATTCG	240
Qy	81	TyrThrIeuSerArgaenGlnThrValValaIgluTyrThriSaapIySaapThrAsp	100
Db	241	TACACGTTGTACACGAGCAGACGCGTGTGTGTGAATACACACGATTAAGAACCGGAC	300
Qy	101	MetPheGlnValaIglArgSerThrGlnSerProIleAspPheValaIThrAspThrIle	120
Db	301	ATGTTTACAGGTGGGACGGTCAACAGAAAGGCCCATTTGATCTTGAGTTCACAGACCGGTT	360
Qy	121	SerGlySerGlnaenThrAspGlnaIaGlnIleThrcInSerThrIleSerArgPheIa	140
Db	361	TCCGGCGGTCAAGACGAAGAT--GCCAGATACACAGAGCACCATCTTAAGTTCCGCA	417
Qy	141	CysArgIleValaCysaapArgaenGlnProTyrThrIlaArgIlePheIaIaIglPhe	160
Db	418	TGCAGAGATCGGTGTGACAGGAAGAGCCCATATATACACACGATATATCGCGCAGAGATTG	477
Qy	161	AspSerSerIySaenIlePheIeuGlnuIyalaIalysTrpIySaenProAspGly	180
Db	478	GATTCCTTCCAAAAATATCTTTCTTGAGAGAGAAACAGCAAAATATGAATAAACCTCGATGGA	537
Qy	181	HisMetAspGlyLeuThrThrAsnGlyValIleuValMetHisProArgGlyGlyPheThr	200
Db	538	CACATGATGATGACTCACTACCAATAGTGTCTCTAATGATGACCCGCAAGAGGCTTCAAC	597
Qy	201	GluGlnSerGlnProGlyValaITrpArgIuIleSerValCysGlyAspValaTyrThrIeu	220
Db	598	GAGGATATCCACGCTGAGAGTCTGAGAGAGATCTCTGTCGTGGGAGATGTATCACTTGG	657
Qy	221	ArgGluThrArgSerIaIaGlnGlnaIyIyIyIeuValGlnSerGluThrAsnValIeu	240
Db	658	CGAAGACACAGGTGGGCCACAGACAGGGGAAAGCTGTGTGAAGATGTAGAACCAAGTCTGTG	717
Qy	241	GlnaapGlySerIeuIleAspIleuCysGlyValaThrIleuLeuTrpArgThrIaIaAspGly	260
Db	718	CMAAGCGCTCTCTCATTTAGCTGTGTGTGGGCCACTCTCTCTGTGAAGACCGCAGATGGC	777
Qy	261	LeuPheHisTrpProthrgInuIyHisIleGluIaIeuArgGlnuIleAsnIaIaIa	280
Db	778	CTTTTTCACGCTCTCACTCAGAAACACATAGAAACCTTCGGCAGAGATCAATGACAGCC	837
Qy	281	ArgProGlnCysProValaGlyLeuIaenThrIeuIaIaPheProSerIleAsnArgIyGlu	300
Db	838	CGAACCCACGTGCCCGGTGGGCTTTAACACCTCTGCCCTTCCACAGATCAACCGAAGGAA	897
Qy	301	ValaIaGluGluIyIyGlnProTrrIaIaTyrIleuSerCysGlyHisIyValaHisGlyTyrHis	320
Db	898	GTGGTGGAAAGAAAGACAGCCTCTGGGCAATCTGTAGCTGGCCCATATGTCAGAGGCTAACAC	957
Qy	321	AsnTrpGlyHisArgSerAspThrGluIaIaAsnGluArgIuCysProMetCysArgThr	340
Db	958	AGCTGGGGCCATCGGAGGAGACGGGGAAGCAAGAGAGGAGTGTCCATGTGACAGACT	1017
Qy	341	ValaGlyProTyrValaIProleuTrpIeuGlyCysGluIaIaGlyPheTyrValaAspIaIy	360
Db	1018	GTGGGCGCCCTTACGTCCCTCTCTGTGGCTGGGCTGTGAGCAGGATTTTATGTGAGACGGGGA	1077
Qy	361	ProProthrhIaIaIaPheThrProCysGlyHisIyValaCysSerGlyIySerAspIaIySer	380
Db	1078	CCCCCAACTACGCGTTTACACCCCTGTGGGGGACGTCGTGTTCAGAAAGTCTGCCAAGTAC	1137
Qy	381	TrpSerGlnIleProleuProHisGlyThrHisIaIaPheHisIaIaIaCysProPheCys	400
Db	1138	TGGTTCGAGATCCACATGCCCCACAGAACGACGCTTATGTGCGGCTGTCCGTTCTGC	1197
Qy	401	AlaThrGlnIeuValaIglGluGlnaenCysIleIyIeuIlePheGlnGlyProIleAsp	420
Db	1198	GCCACGCGAGCTGTGTGTGTGAACAGACATGCATCAATATGATTTTCCAAAGTCCAGTGAC	1257

DEFINITION Xenopus laevis MGC80329 protein, mRNA (cDNA clone MGC:80329
IMAGE:5074048), complete cds.
ACCESSION BC072891
VERSION BC072891.1 GI:49119163
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 1646)
Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
REFERENCE 2 (bases 1 to 1646)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Shapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaeetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramsen, R.D., Muliyil, S.J., Bosak, S.A., McGowan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W.,
Villalon, D.R., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fenche, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shchepenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)
JOURNAL 12477932
PUBMED 3 (bases 1 to 1646)
REFERENCE Klein, S. and Gerhard, D.S.
AUTHORS Direct Submission
JOURNAL Submitted (07-JUN-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
REMARK NIH-MGC Project
COMMENT Contact: XGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Igor David
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prihbu,
Parvaneh Saeedi, Jr Santos, Angeliq Scherch, Ursula Skalska,
Duane Smalins, Jeff Stott, Miranda Tsai, George Yang, Jacqui
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAX Plate: 152 Row: h Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
Location/Qualifiers
1..1646

gene
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:80329 IMAGE:5074048"
/cbase_type="Ovary, Xenopus"
/clone_lib="NICHD_XGC_OV1"
/lab_host="DH108"
/note="Vector: pCMV-Sport6"
1..1646
/gene="MGC80329"
/db_xref="GeneID:443954"
336..1595
/gene="MGC80329"
/codon_start=1
/product="MGC80329 protein"
/protein_id="AAH72891.1"
/db_xref="GI:49119164"
/db_xref="GeneID:443954"
/translation="MFSPEERHCAKSPVYXGELVILGNGCLPENGDRGRKSRPA
LFRKPKSNVRASSVHISTPOAKAISRGCHSIFLISRSQTVVETHTDKDPMF
OIGRSTSPIDFVTVGVGONDETOITOSTISRAICVDRNPYARLPAAGD
SSKNIFLGKAKAKWKNPDGMDGLTTNGVLVNHKPGFTDESFGVREISVGGDYT
LHRTSAQRGKLVERTNIIIDQGSIVDCATILMRTADGLHTPTQGHIAKREI
NAAPQCPVGLNTLAPSPSINRQDVDEKQPMVILKGVHGHINMGHRSDDTEANREC
PMKRTIGPVVPLMLGCEAGPYVDVPTPAFPCGHCSEKSAKWSOILPLPGTIAF
HAACPFCALQITGSEBSCVILIFQSPVD"

ORIGIN

Alignment Scores:
Pred. No.: 6,84e-155 Length: 1646
Score: 2138.50 Matches: 383
Percent Similarity: 96.9% Conservative: 24
Best Local Similarity: 91.2% Mismatches: 12
Query Match: 93.4% Indels: 1
DB: Gaps: 1

US-10-041-030-4 (1-420) x BC072891 (1-1646)

QY 1 MetPheSerProGlyngIngluInuHsCyBaLaProAsnLysGluProValLysTyrGly 20
DB |||||CCCCCGAGTCAAGGAGAACATTGGGCCCCGATAGAGGCCGTTGAATACGGA 335
QY 21 GluLeuValValLeuGlyTyrAaGlyAlaLeuProAsnGlyAspArgGlyArgTyrGly 40
DB |||||GAGCTGGTGTGGGGGTACACAGGCTGTCTCCAAATGGGATGAGGAAGCCGGA 455
QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysPheSerThrValHis 60
DB |||||AGCAGATTGCACCTTCAAAACCCCAATCTMAAGGGGTGAGACCCAGCTGTTCAT 515
QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLeuGlyngInuHisSerIleSer 80
DB |||||GTGATTTCAACCCCGAGGCTTCCAGGCAATGACAGCAGAGCCCAACACACATCTCA 575
QY 81 TyrThrLeuSerArgAaGlnInuThrValValAlaGluTyrThrHisAspLysAspThrAsp 100
DB |||||TTCAATTATCCAGAGAGTGAAGCTGTTGGTGGTAATACCAATGACCAAGACAGAT 635
QY 101 MetPheGlnValAlaGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
DB |||||ATCTTTCAAGTTGGAAGGTCAACGAAGCCCATTTGATGTAATCTGATACGGGT 695
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB |||||TCTGGCAACGAGAAC--GATGAACACAGATACCAAAATTCAAATTCAGATTGCA 752
QY 141 CysArgIleValCysAspArgAaGlnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
DB |||||TGTAGATGATCTGCATAGAACCCGCGTACAGAGTATATTTGTCAGAGATTT 812
QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTyrLysAsnProAspGly 180
DB |||||GACTCTCCAAAATATATTTCTTGGTGAAGAACGGCCAAAGTGAAGAAATCTGACGGT 872

Alignment Scores:

Pred. No.: 4,05e-154 Length: 1936
 Score: 2129.50 Matches: 384
 Percent Similarity: 96.7% Conservative: 22
 Best Local Similarity: 91.4% Mismatches: 13
 Query Match: 93.0% Indels: 1
 Gaps: 1

US-10-041-030-4 (1-420) x BC063200 (1-1936)

```

QY 1 MetPheSerProGlyngluGluHisCyseAlaProAsnlysgluProVallysrTgLy 20
DB 365 ATGTTTCCCCAGTCAGAGAGACACTGGCCCCCGTAAGAGCCCGCAATATAGGA 424
QY 21 GluLeuValIleuGlyTyrAsnGlyAlaLeuProAsnlysaPaArgLy 40
DB 425 GAGCTGGTGGTGGGGGTACACGGCTGTCTCCCAATGGGGATCGAGAAAGCGGAAA 484
QY 41 SerArgPheAlaLeuTyrLyAsArgProLyAlaAsnGlyVallysrProSerThrValHis 60
DB 485 AGCAGATTTCACCTTTAAAGCCCTTAATCCATGGGGTGAAGCCAGCAGTTCAT 544
QY 61 ValIleSerThrProGlnAlaSerLyseAlaIleSerCyseLysglyGlnHisSerIleSer 80
DB 545 GTGATTTCACACCCAGAGCTTCACAGGCAATCGACAGAGAGGCCAACACAGCATCTCA 604
QY 81 TyrThrIleuSerArgAsnGlnThrValValIleGluTyrThrHisAaPlysaPThrAsp 100
DB 605 TTCCATTATTCACAAATGACACTGGTGTGTGGTAATACCCATGACAGGACACAGAT 664
QY 101 MetPheGlnValGlyArgSerThrGlySerProIleAsPheValIleThrAspThrIle 120
DB 665 ATGTTTCAAGTGGAAAGGTCAACAGAAAGTCCCATTACTTGTAGTACTGACACGGTT 724
QY 121 SerGlySerGlnAsnThrAspGlyValIleGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 725 TCTGGCAACACAGAAAT---GATGAACAACAGATAACAAAGTCAATCTCAGATTGGC 781
QY 141 CyseArgIleValCyseAspArgAsnGluProTyrThrIleArgIlePheAlaIleGlyPhe 160
DB 782 TGCAGATGATCTGTGACAGAAACCTCCCGTACACCCCTGAATATTTGTCGACGAGATT 841
QY 161 AspSerSerLyseAsnIlePheLeuGlyGluValIleAlaIleValIleValIleAsnProAspGly 180
DB 842 GACTCCCTCCAAATAATATTTCTTGTGGAGAAAGCGCCAAAGTGAAGATCCGATGCT 901
QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValIleMetHisProArgGlyGlyPheThr 200
DB 902 CATATGATGAGACTAACCTAACCAATGTGTCTTGTGATGATCCGAAGGAGGGGTTACT 961
QY 201 GluGluSerGlnProGlyValIleTPArgGluIleSerValCyseGlyAspVallysrThrLeu 220
DB 962 GAAGAAATCCAGGCTGAGATGAGCGGGAATTTCACTGTCTGAGATGTATATACATTA 1021
QY 221 ArgGluThrArgSerAlaGlnIleArgGlyValIleValIleGluSerGluThrAsnValIleu 240
DB 1022 AGAGAAACCGATTTGCTACGCAAGAGGCAAACTTGTGGAAATGAAACCAATATCTG 1081
QY 241 GlnAspGlySerLeuIleAspLeuCyseGlyValIleThrLeuLeuTPArgThrAlaAspGly 260
DB 1082 CAGATGTTCTCTTGTGACTTATGTGTGCCACCTTTCTTGGCAACGCGATGCT 1141
QY 261 LeuPheHisThrProThrGlnIleValHisIleGluAlaLeuArgGlnIleValIleAsnAla 280
DB 1142 TTGCTCATACACGACTCAAAAACACATTGACCTTGAGACCGGAATTAATAGCTGCC 1201
QY 281 ArgProGlnCyseProValIleValAsnThrIleuAlaPheProSerIleAsnArgLyGlu 300
DB 1202 AGGCTCAGTGCCTGTGTGATTAACAATTACCTTTCCAGATACAAACCGAAGAT 1261
QY 301 ValIleGluGluValysGlnProTPAlaIleTyrLeuSerCyseGlyHisValIleGlyTyrHis 320
  
```

```

DB 1262 GTCTAGAAAGAGAAACAGCCTTGGGCTTACCTGACCTGTGGTCAATGTCAGGCTATCAT 1321
QY 321 AsnTPGlyHisArgSerAspThrGluAlaAsnGluArgGluCyseProMetCyseArgThr 340
DB 1322 AACTGGGGTATGTCAGTACACAGAAAGTCAATGAACGCGAGTCCCATGTGTAGAAC 1381
QY 341 ValGlyProTyrValIleProLeuTPLeuGlyCyseGluAlaIleGlyPheTyrValIleAsp 360
DB 1382 GTTGGTCCCTATGCTGCTTATGCTGTGGTGGTGAAGACAGATTTCAAGTAAATGCTGGA 1441
QY 361 ProProThrHisAlaPheThrProCyseGlyHisValIleCyseSerGluValysSerAlaIleTyr 380
DB 1442 CCTTCCTCACTCACCTTTCAAACTTGTGACAGCTCAGCTCAGAAAAATCTGCCAATAT 1501
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIleCyseProPheCyse 400
DB 1502 TGGTTCAGATTTCACCTGACCAATGGTACACAGCGCTTTCATCTGCTTCTTCTGC 1561
QY 401 AlaThrGlnLeuValIleGluGlnAsnCyseIleValIleLeuIlePheGlnGlyProIleAsp 420
DB 1562 GCATATCAATTAATGACTGGAAGAAAGAACTGTGTCAAACTTATTTCCAGGGCCCTGTTGAC 1621
  
```

RESULT 12
 BD190327
 LOCUS 1717 bp DNA linear PAT 17-JUL-2003
 DEFINITION BklI phosphorylation related gene.
 ACCESSION BD190327
 VERSION BD190327.1 GI:33000066
 KEYWORDS WO 03008589-A/67.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS Matsuzaki, O., Matsumoto, A., Nagano, Y. and Suzuki, N.
 TITLE BklI phosphorylation related gene
 JOURNAL Patent: WO 03008589-A 67 30-JAN-2003;
 ASAHII KASEI CORP, OSAMU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI

COMMENT OS Mus musculus (mouse)
FN WO 03008589-A/67

PD 30-JAN-2003
 PR 15-JUL-2002 WO 2002JP007174
 PR 18-JUL-2001 JP 01P 218204, 31-AUG-2001 JP 01P 263450 PR
 21-JAN-2002 JP 02P 012176
 PI OSAMU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI PC
 C12N15/54, C12N9/12, G01N33/50, G01N33/15, C07K16/40, A61K39/385, PC
 A61K31/711

CC BklI phosphorylation related gene
 FH Key location/Qualifiers
 FT CDS (218) . (1474).

FEATURES
 source location/Qualifiers
 1..1717
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"

ORIGIN

Alignment Scores:
 Pred. No.: 8.55e-153 Length: 1717
 Score: 2111.50 Matches: 388
 Percent Similarity: 95.2% Conservative: 12
 Best Local Similarity: 92.4% Mismatches: 19
 Query Match: 92.2% Indels: 1
 Gaps: 1

US-10-041-030-4 (1-420) x BD190327 (1-1717)

```

QY 1 MetPheSerProGlyngluGluHisCyseAlaProAsnlysgluProVallysrTgLy 20
DB 218 ATGTTTCCCCGAGGAGAGAAACCCAGCGCCCCCAACAAAGAGCCGGTGAATACAGG 277
  
```

QY 21 G|U|L|U|V|A|L|L|U|G|L|Y|T|Y|R|A|N|G|L|Y|A|L|L|U|P|R|O|A|N|G|L|Y|A|S|P|A|R|G|L|Y|A|R|G|L|Y|S 40
 |||||
 DB 278 G|A|C|T|G|T|G|T|C|T|G|G|G|T|C|A|A|T|G|T|G|C|T|T|A|C|T|A|T|G|T|G|A|C|A|G|G|G|C|A|A|G|A|A 337
 |||||
 QY 41 S|E|R|A|R|G|P|H|E|A|L|E|U|T|Y|R|L|Y|S|A|R|G|P|R|O|L|Y|A|I|A|S|N|G|L|Y|A|L|L|Y|S|P|R|O|S|E|R|T|H|A|L|H|S 60
 |||||
 DB 338 A|G|C|A|G|A|T|T|A|C|C|T|A|T|A|A|G|G|G|A|C|C|T|A|G|C|A|G|G|T|G|T|C|A|A|C|C|C|A|C|A|A|T|C|C|A|C 397
 |||||
 QY 61 V|A|L|L|S|E|R|T|H|P|R|O|G|L|A|I|S|E|R|L|Y|S|A|I|L|E|S|E|R|C|Y|L|S|G|L|Y|G|L|N|H|I|S|E|R|L|E|S|E|R 80
 |||||
 DB 398 A|T|G|T|C|C|C|A|C|A|C|A|G|G|G|C|C|A|G|G|C|C|A|T|G|C|C|C|A|G|A|G|A|C|A|C|A|T|C|C|A|C|A|T|A|T|C|G 457
 |||||
 QY 81 T|Y|R|T|H|L|E|U|S|E|R|G|A|N|G|L|I|N|T|H|A|L|V|A|L|G|L|T|Y|R|H|H|I|S|A|P|L|Y|S|A|R|G|P|R|A|R|P 100
 |||||
 DB 458 T|A|C|A|C|G|T|G|T|C|A|G|G|C|C|A|G|C|A|G|G|T|G|T|G|A|G|A|C|A|C|A|C|A|T|A|A|A|G|C|A|C|A|G|A|C 517
 |||||
 QY 101 M|E|T|P|H|E|G|L|I|N|V|A|L|Y|R|S|E|R|T|H|G|L|U|S|E|R|P|R|O|L|E|S|P|H|E|V|A|L|T|H|A|R|A|P|R|H|L|E 120
 |||||
 DB 518 A|T|G|T|T|C|A|G|T|G|G|G|A|G|T|C|A|C|A|G|A|A|G|C|C|A|T|T|G|A|C|T|T|G|G|T|G|C|A|C|A|G|A|C|G|G|T 577
 |||||
 QY 121 S|E|R|G|L|Y|S|E|R|G|A|N|T|H|A|R|G|L|U|A|G|L|I|L|E|T|H|G|L|I|N|S|E|R|T|H|L|E|S|E|R|A|R|G|P|H|E|A|L 140
 |||||
 DB 578 T|C|C|G|G|G|G|T|C|A|G|A|C|G|A|A|G|A|T|---G|C|C|A|G|A|T|C|A|C|A|G|A|C|A|C|A|T|C|T|A|G|G|T|G|C|A 634
 |||||
 QY 141 C|Y|E|A|R|I|L|E|V|A|L|Y|S|A|P|A|R|G|A|N|G|L|I|N|P|R|O|T|Y|R|T|H|A|R|G|L|E|P|H|E|A|I|A|L|Y|P|H|E 160
 |||||
 DB 635 T|G|C|A|G|A|T|G|T|G|T|G|A|C|A|G|A|A|C|G|A|C|A|C|A|G|C|A|G|C|A|T|T|C|G|G|G|C|A|G|A|T|T|C 694
 |||||
 QY 161 A|S|P|S|E|R|L|Y|S|A|N|I|L|E|P|H|E|U|G|L|Y|L|U|Y|A|A|L|A|L|Y|T|R|P|L|Y|S|A|N|P|R|O|A|R|P|G|L|Y 180
 |||||
 DB 695 G|A|T|T|T|C|A|A|A|A|A|T|C|T|T|T|T|G|A|G|A|A|A|G|C|A|A|A|A|T|G|A|A|A|A|C|C|T|G|A|T|G|A 754
 |||||
 QY 181 H|I|S|E|R|A|S|P|G|L|Y|L|E|U|T|H|T|H|A|N|G|L|Y|A|L|L|E|U|V|A|L|W|E|T|H|I|S|P|R|O|A|R|G|L|Y|G|L|Y|P|H|E|T|H 200
 |||||
 DB 755 C|A|T|G|A|T|G|A|T|G|A|C|T|A|C|A|C|A|G|G|G|T|C|T|G|T|G|A|T|G|A|C|C|C|G|C|A|G|A|G|G|C|T|C|A|C 814
 |||||
 QY 201 G|L|U|G|L|S|E|R|G|I|N|P|R|O|G|L|Y|A|L|T|P|R|A|R|G|L|I|L|E|S|E|R|V|A|L|C|Y|S|G|L|Y|A|S|P|A|R|V|A|L|Y|R|T|H|L|E|U 220
 |||||
 DB 815 G|A|G|A|A|T|C|C|C|A|G|C|T|G|A|G|T|G|A|G|A|G|A|T|C|T|G|T|G|G|G|A|G|A|T|A|C|A|C|C|T|T|G 874
 |||||
 QY 221 A|R|G|L|U|T|H|A|R|G|S|E|R|A|G|I|N|G|L|I|N|A|R|G|L|Y|L|U|Y|L|E|U|V|A|L|G|L|S|E|R|G|I|N|T|H|A|N|V|A|L|L|E|U 240
 |||||
 DB 875 C|G|A|G|A|A|C|C|A|G|T|C|G|G|C|C|A|G|C|A|G|A|G|G|G|A|A|G|C|T|G|T|G|A|A|A|G|T|G|A|C|A|C|A|G|C|T|C|G 934
 |||||
 QY 241 G|I|N|A|S|P|G|L|Y|S|E|R|L|E|U|L|E|A|S|P|L|E|U|C|Y|S|G|L|Y|A|L|T|H|L|E|U|E|U|T|P|R|A|R|G|T|H|A|L|A|S|P|G|L|Y 260
 |||||
 DB 935 C|A|G|A|C|G|G|T|C|C|C|A|T|T|G|A|C|T|G|T|G|G|G|C|C|A|C|T|C|T|C|T|G|A|G|A|A|C|G|C|A|A|T|G|C|A|G|C 994
 |||||
 QY 261 L|E|U|P|H|E|I|S|T|H|P|R|O|T|H|G|L|I|N|Y|H|I|L|E|G|L|U|A|L|E|U|A|R|G|L|I|N|L|I|L|E|A|N|A|I|A|L|A 280
 |||||
 DB 995 C|Y|T|T|T|C|A|C|G|C|T|C|T|C|T|C|A|G|A|G|C|A|C|A|T|G|A|A|G|C|C|C|C|G|C|A|G|G|A|G|A|T|G|C|A|G|C|C 1054
 |||||
 QY 281 A|R|G|P|R|O|G|L|Y|S|P|R|O|V|A|L|G|L|Y|L|E|U|A|N|T|H|L|E|U|A|L|A|P|H|E|R|O|S|E|R|L|E|A|N|A|R|G|L|Y|G|L|U 300
 |||||
 DB 1055 C|A|G|C|C|C|A|G|G|C|C|C|G|G|G|C|C|T|T|A|C|A|C|C|C|T|G|C|T|T|C|C|C|A|G|C|A|T|C|A|C|C|G|A|G|A|A 1114
 |||||
 QY 301 V|A|L|V|A|L|G|L|U|L|Y|S|G|I|N|P|R|O|T|P|R|A|L|Y|R|L|E|U|S|E|R|C|Y|S|G|L|Y|H|I|S|V|A|L|H|I|S|G|L|Y|T|Y|R|H|S 320
 |||||
 DB 1115 G|T|G|T|G|A|A|G|A|G|A|G|A|G|C|C|T|G|G|G|A|T|C|T|G|A|G|T|G|G|G|C|A|T|G|G|C|A|T|G|G|C|T|A|C|A|C 1174
 |||||
 QY 321 A|S|E|R|T|G|L|Y|H|I|S|A|R|G|S|E|R|A|R|P|R|H|G|L|U|A|A|S|N|G|L|U|R|G|L|U|C|Y|S|P|R|O|T|H|E|C|Y|S|A|R|G|T|H| 340
 |||||
 DB 1175 C|A|C|T|G|G|G|G|C|A|T|G|G|G|C|A|G|C|A|G|G|A|G|C|C|A|G|A|G|G|G|A|G|G|T|C|C|C|A|T|G|C|A|G|C|A|C|T 1234
 |||||
 QY 341 V|A|I|G|L|Y|P|R|O|T|Y|R|A|L|P|R|O|L|E|U|T|P|R|E|U|G|L|Y|C|Y|S|G|L|U|A|L|A|P|H|E|R|Y|V|A|L|A|S|P|A|L|G|L|Y 360
 |||||
 DB 1235 G|T|G|G|G|C|C|C|A|G|C|C|C|T|C|T|G|G|C|T|G|G|C|T|G|G|G|C|A|G|A|G|A|T|T|A|T|G|T|G|C|A|N|N|N|N|N 1294
 |||||
 QY 361 P|R|O|P|R|O|T|H|H|I|S|A|L|A|P|H|E|R|P|R|O|C|Y|S|G|L|Y|H|I|S|V|A|L|C|Y|S|S|E|R|G|I|U|Y|S|E|R|A|L|Y|E|R|Y 380
 |||||
 DB 1295 N|N|N|N|N|N|N|C|A|T|G|T|T|T|C|A|C|C|C|C|T|G|G|G|C|A|C|G|T|G|T|C|A|G|A|G|A|G|T|G|C|C|A|G|A|T|C 1354
 |||||
 QY 381 T|R|P|S|E|R|G|I|N|L|E|P|R|O|L|E|U|P|R|O|H|I|S|G|L|Y|T|H|H|I|S|A|L|A|P|H|E|I|S|A|L|A|C|Y|S|P|R|O|P|H|E|C|Y|S 400
 |||||

DB 1355 T|G|T|T|C|C|A|A|T|T|C|C|A|T|G|C|C|C|C|A|G|A|C|A|G|C|G|T|T|C|A|G|C|C|G|C|T|G|C|G|T|C|T|G 1414
 |||||
 QY 401 A|A|T|H|G|L|U|V|A|L|G|L|U|G|L|U|A|N|C|Y|S|I|L|E|S|E|R|L|E|P|H|E|G|I|N|G|L|Y|P|R|O|L|E|A|R|P 420
 |||||
 DB 1415 G|C|C|A|C|C|A|C|T|G|G|T|G|T|G|A|C|A|G|A|C|A|T|C|A|T|A|A|T|G|A|T|T|T|C|A|A|G|T|C|A|G|T|G|A|C 1474
 |||||
 RESULT 13
 AF302504
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridea; Muridae; Murinae; Mus.
 1 (bases 1 to 1734)
 Resch, K., Jockusch, H. and Schmitt-John, T.
 Assignment of homologous genes, Pell1/Pell1 and Pell2/Pell2, for
 the Pelle adaptor protein Pellino to mouse chromosomes 11 and 14
 and human chromosomes 2p13.3 and 14q21, respectively, by physical
 and radiation hybrid mapping
 Cytogenet. Cell Genet. 92 (1-2), 172-174 (2001)
 11306823
 JOURNAL
 PUBMED
 2 (bases 1 to 1734)
 Resch, K., Jockusch, H. and Schmitt-John, T.
 Direct Submision
 Submitted (05-SEP-2000) Developmental Biology and Molecular
 Pathology, NRW 33615, Germany
 Bielefeld, University of Bielefeld, Universitaetsstrasse 25,
 Bielefeld, Location/Qualifiers
 1. 1734
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 1. 1734
 /gene="Pell2"
 218..1477
 /gene="Pell2"
 /note="similar to Drosophila melanogaster Pellino"
 /codon_start=1
 /product="pellino 2"
 /protein_id="AA015392.1"
 /db_xref="GI:10242357"
 /translation="MPSPOEERSPANPKRPKYRELIVLGYNGALPENGRRKKGRFT
 LKRYTASGVKPSLHMVSTPQASKAISRGHSISTYLSQTVVETHTDQDMF
 QVGRSTESPIDIVVDVTSVGGQNEAQLQSTISRAKRIVDNRNPYARIIPAAPD
 SSKNIFLGEKAKMKNDGMDGLITNGQIVLNMHPDGFPEESQPGVWRISVCGDVT
 LKRTSAOQGLVSEETNVLQDGLSIDCGATLMTKTDGFLHAPYRKHIEAROEI
 NBARPCPYGLNTLAPEPSINREKREVERKPMAYLSCGHVGYHMGHRSDETEAREC
 ENQMTVGPVPLMIGCEAFYLYXXXXXHVLPFGCHVSEKSAKTYSLPLPHGTAF
 HAACPFCATQLVGBQNCIKLIFQGPVD"
 ORIGIN
 Alignment Scores:
 Pred. No.: 8, 64e-153 Length: 1734
 Score: 2111.50 Matches: 388
 Percent Similarity: 95.2% Conservative: 12
 Best Local Similarity: 92.4% Mismatches: 19
 Query Match: 92.2% Indels: 1
 DB: 9 Gaps: 1
 US-10-041-030-4 (1-420) * AF302504 (1-1734)
 QY 1 M|E|T|P|H|E|S|E|R|P|R|O|G|L|Y|G|L|U|G|L|U|H|I|S|C|Y|A|L|A|P|R|O|A|N|L|Y|S|G|L|U|P|R|O|V|A|L|Y|R|Y|G|L|Y 20
 |||||
 DB 218 A|T|G|T|T|T|C|C|C|G|G|G|C|A|G|A|G|A|A|C|C|C|C|C|A|C|A|G|A|G|C|G|G|T|G|A|A|A|T|A|C|A|G 277
 |||||
 QY 21 G|U|L|U|V|A|L|L|U|G|L|Y|T|Y|R|A|N|G|L|Y|A|L|L|U|P|R|O|A|N|G|L|Y|A|S|P|A|R|G|L|Y|A|R|G|L|Y|S 40
 |||||


```

|||||.....
278 GACGTGTCGTCGGGGTCAATAGTGTCTTACTATGTCAGACGGGCAAGAAGAA 337
Qy |||||
41 SerArgPheAlaLeuTyrIlySarGProLyAlaAsnIlyValIyPProSerThrValHis 60
Db |||||
338 AGCAGATTATCCCTTATTAAGCGGACCTACGCGCCGAGTGTCTCAAAACCACACATCAC 397
Qy |||||
61 ValIleSerThrProGlnAlaSerIyAlaIleSerCysIlyGlnIleSerIleSer 80
Db |||||
398 ATGTCCTCCACACACAGCGGCTCCAAAGCCATCAGCTCCAGAGAACATCACAGATTCG 457
Qy |||||
81 TyrThrIleuSerArgaenGlnThrValValIyIlyThrHisIlyIlyIlyIlyIlyIly 100
Db |||||
458 TACACGTTGTACAGAGCCAGACAGCGTAGTGTGAGTACACACGATTAAGACACAGAC 517
Qy |||||
101 MetPheGlnValIyIySarSerThrGlnIyPProIleAspPheValIlyThraAspThrIle 120
Db |||||
518 ATGTTTCAGGTGGGACAGTCAAGAAAGCCCATTAATTCGTGTGACACAGACGCGTT 577
Qy |||||
121 SerGlySerGlnAsnThraAspGlnIyIyGlnIleThrGlnSerThrIleSerArgPheAla 140
Db |||||
578 TCCGGGGCGTCAGAAACGAAGAT---GCCAGATCACACAGACCATCTTACGTTCCGA 634
Qy |||||
141 CysArgIleValIyCysAspArgaenGlnProTyrThraIlyIlyIlyPheAlaIlyIlyPhe 160
Db |||||
635 TGCAGAGTCGTGTGTGACAGAAACGACCATACACAGCATATTCCGCGCGAGATTC 694
Qy |||||
161 ArgSerSerIyAsnIlePheIleuGlnIlyIyAlaIlyIlyIlyIlyIlyIlyIlyIlyIly 180
Db |||||
695 GATTTCTCCAAATAATCTTTCTTGGAGAAAGACCAAAATGGAAAAACCTCGAATGGA 754
Qy |||||
181 HiMetAspGlyLeuThrThraAsnGlyValIleuValIleMetIlyProArgIlyGlyPheThr 200
Db |||||
755 CACATGATGATGATCTACATCAACAGGCTGTCTGTGTATGCAACCCGCAAGAGGCTTACC 814
Qy |||||
201 GlnGlnSerGlnProGlyValIlyThraGlnIlyIleSerValIyCysIlyAspValIlyThrIleu 220
Db |||||
815 GAGGAATCCAGCGCTGAGTGTGAGAGAGATCTCTGTGTGGGAGTATACACCTTG 874
Qy |||||
221 ArgGlnThrArgSerAlaGlnIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 240
Db |||||
875 CGAGAGCCAGGTCGCGCCAGCAAGGGAAGCTGTGTGAAGATGAGCAACAGCTCTG 934
Qy |||||
241 GlnAspGlySerLeuIleAspLeuIyCysGlyValIleThrLeuIyThraGlyThraIlyAspGly 260
Db |||||
935 CAAAGCGGCTCTCTCATTTGACTGTGTGGGCGCATCTCTCTGGAACACCGCATGCGC 994
Qy |||||
261 LeuPheIleThrProThrglnIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 280
Db |||||
995 CTTTTCACGCTCTACTACAGAACATAGAACCTCCGCGAGAGATCAATGACAGCC 1054
Qy |||||
281 ArgProGlnIySarProValIyIlyAsnThrIleuAlaPheProSerIleAsnArgIyGln 300
Db |||||
1055 CGAGCCCAAGGCGCCGCGGCTTAAACCTTGGCTTCCCGACATCAACCGGAAGAA 1114
Qy |||||
301 ValIyGlnIlyIyGlnProThraIyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 320
Db |||||
1115 GTGTGTGAAGAAGAGCGCCCTGCGCATCTGAGCTGCGGCGCATGCGACGCGTTACAC 1174
Qy |||||
321 AsnIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 340
Db |||||
1175 CACTGGGGCGCATCGAGACGACACGGAAGCCAAAGAGGGAGTGTCCCATGTGATGACT 1234
Qy |||||
341 ValIyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 360
Db |||||
1235 GTGGGCGCCCAAGCGCTCTCTGTGCTGTGAGTGTGAGCAGAGATTTTATGCGANNNNNN 1294
Qy |||||
361 ProProThrhIlyAlaPheThrProCysGlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 380
Db |||||
1295 NNNNNNNCTATGTTTTCACCCCTGCGGAGCATCTGTTCAGAGAGTGTGCCAATAC 1354
Qy |||||
381 TrpSerGlnIlyProLeuProHleGlyThrhIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 400

```

```

Db 1355 TGTCCAGATCCCATGCCGCCAGGACGACGCGCTTTCACGCGCTGTCCCTTCG 1414
Qy 401 AlaThrglnIyValIyGlnIyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 420
Db 1415 GCCACGAGCTGTGTGTGAAGAACTGATCAATATGATTTTCCAAAGTCCAGTGGAC 1474

RESULT 14
LOCUS BC075973
DEFINITION Danto rerio pellino homolog 2 (Drosophila), mRNA (cDNA clone
ACCESSION BC075973.1 GI:49903830
VERSION BC075973.1
KEYWORDS MGC.
SOURCE Danto rerio (zebrafish)
ORGANISM Danto rerio

REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Musina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stepleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheltz, T.E., Brownstein, M.J., Usdin, T.B., Tashy, T.S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Adams, R.D., Muliyil, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S.,
Sanchez, J., Whitting, M., Madan, A., Young, A.C., Shcherbakov, Y.,
Boutard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Small, D.B.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2474)

JOURNAL NIH MGC Project
PUBMED Direct Submission
REFERENCE Submitted (06-JUL-2004) National Institutes of Health, Mammalian
AUTHORS Gene Collection (MGC), Bethesda, MD 20892-2590, USA
JOURNAL NIH-MGC Project URL: http://mgc.ncl.nih.gov
REMARK Contact: MGC help desk
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 179 Row: m Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 49903830.
location/Qualifiers
1..2474
/organism="Danto rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:92268 IMAGE:7053197"
FEATURES
source

```


Alignment Scores:

Pred. No.:	5,34e-138	Length:	1257
Score:	1917.00	Matches:	342
Percent Similarity:	90.5%	Conservative:	38
Best Local Similarity:	81.4%	Mismatches:	38
Query Match:	83.7%	Indels:	2
		Gaps:	1

US-10-041-030-4 (1-420) x AR483566 (1-1257)

```

QY      1 MetPheSerProGlnGlnGluHISCyaaLaProAenlySGluProValIlyeTyGly 20
DB      1 ATGTTTCTCTGATCAAGAAATCAT-----CCATCTAAGACACAGATAAATATGTT 54
QY      21 GluLeuValIleuGlyTyraenGlyAlaLeuProbenGlyAaPaAGlyAaAGAGly 40
DB      55 GAACTCATTTGCTTAAGATTAATAGATCTCTCCAAACGGTATAGAGAGAGAGAGAAA 114
QY      41 SerArgPheAlaLeuTyTyraenGlyAaPaAGlyAaAGlyAaAGAGlyAaAGAGly 60
DB      115 AGTAGTTTCTCTGATCAAGAAATCAT-----CCATCTAAGACACAGATAAATATGTT 174
QY      61 ValIleSerThrProGlnAlaSerIyAlaIleSerCySlySGlyGlnHISerIleSer 80
DB      175 ATTGCTTGTACTCTCTGAGCTGCAAGAGCAATAGCAAAAGACACAGATAGCATATCA 234
QY      81 TyrThrIleuSerArgAenGlnThrValIValIValIValIValIValIValIValIVal 100
DB      235 TATACCTTATCTCGGCGCCAGACCTGTGTGATGATTAATCTCATGACACCAACCGAT 294
QY      101 MetPheGlnValGlyArgSerThrGluSerProIleAaPheValIValIThrAaPThrIle 120
DB      295 ATGTTTCAGATTTGGCCGGTGCAGCTGAAGCCCATTTGTTTGTATGTAAGTGAACGGTT 354
QY      121 SerGlySerGlnAenThrAaPheGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB      355 CCTGGAAGTCAAGTAATTTCTGATACACAGTCAAGTCAAAAGCACTATATCAAGATTGCC 414
QY      141 CyaaArgIleValCyaaPaArgAenGluProTyThrIlaArgIlePheAlaIAGlyPhe 160
DB      415 TGGAGAAATCAATAGTGAAGAAATCTCTCCCTTACACACGGATTTATGCTGACGGGTTT 474
QY      161 AspSerSerIyAaenIlePheLeuGlyGluYaaIaIaIySTpIyAaenProAaPGLy 180
DB      475 GACTCATCAAAAACATCTTTCTTGGGGAAGAAGCTGCCAAATGGAAGACATCATGATGA 534
QY      181 HisMetAaPGLyLeuThrAaenGlyValIleuValIleuValIleuValIleuValIleu 200
DB      535 CAGATGATGGCTTGACCACTAATGATGTTCTTGTGATGATCCACCAATGGGTTTACA 594
QY      201 GluGluSerGlnProGlyValITPArgGluIleSerValCySGlyAaPValITyThrLeu 220
DB      595 GAAAGACTCCAAAGCTGGAATATGGAAGAAATATCGGTGTGTGGAAATGTAATTAAGCTTA 654
QY      221 ArgGluThrArgSerAlaGlnGlnAaArgIlyLeuLeuValIleuSerGluThrAaenValIleu 240
DB      655 CGTGAACCAAGATGGCTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
QY      241 GluAaPGLySerIleuIleAaPLeuCySGlyAlaThrIleuLeuITPArgThrIlaaSPGLy 260
DB      715 CAAGATGGCTCGTAAATTTGACCTGTGTGACCAATTTGATGCTGCTGCTGCTGCTGCTG 774
QY      261 LeuPheIleThrProThrGlnIlyuYaaIleGluAlaLeuArgGlnGlnIleAaenAlaIa 280
DB      775 CTTTCCCACTCTCTACCGTGAAGCATTTAGAACCTTTAAGACAGAAATCAATGACAGCA 834
QY      281 ArgProGlnCySPProValGlyLeuAaenThrIleuIlaPheProSerIleAaAaArgLyGlu 300
DB      835 CGACCTGAGTCCCTGTAGGGTTCAACACACTAGCATTTCTTATGATGAAGAGAGAGAGAC 894
QY      301 ValIleGluGlnIlyuYaaIleuITPAlaITyThrIleuSerCySGlyAlaIleuValIleu 320
DB      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

```

DB      895 GTTGTAGATGAAAAACAACCATGGGTATATCTAAACTGGGGCCCATGTACATGCTATCAT 954
QY      321 AaenTrGlyHisArgSerAaPThrGluAlaAenGluAaGlyuCySPrometCyAaArgThr 340
DB      955 AACTGGGGAACAAAGAAAGAGCTGATGGCAAGATCGTAATGTCTTATGTATAGTCT 1014
QY      341 ValGlyProTyValIleuLeuITPLeuGlyCySGlyAlaGlyPheTyValAaPaAGly 360
DB      1015 GTTGTGCCATGTTCTCTGTGTGCTTGGATGTGAAGCTGAATTTATGTGACGCCGGC 1074
QY      361 ProProThrHisAlaPheThrProCySGlyHisValCySerGluYaaSerAlaIyTy 380
DB      1075 CTTCCAAACCATCGTTTACCGCTGTGGCATGTGTTCAGAAAAAGCACTGCTTAT 1134
QY      381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIaCySPProPheCy 400
DB      1135 TGGTCCAGATCCCACTTCTCTCAAGTACTCATTTTCAATGACACGCTGTCTTTGT 1194
QY      401 AlaThrGlnLeuValIleGlnIleAaenCySlyLeuLeuIlePheGlnGlyProIleAaP 420
DB      1195 GCACATCAGTTGGCTGTGAAACAAGGCTACATCAGATTATTTTCAAGGACCTTAAGAC 1254

```

Search completed: March 2, 2006, 07:33:41
 Job time : 6959 sec

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 21:01:56 ; Search time 20 Seconds
(without alignments)
418.898 Million cell updates/sec

Title: US-10-041-030-4

Perfect score: 2290
Sequence: 1 MESPGEHCAHPKPKYK.....ATQVGEQNCIKIFQGPID 420

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 133702 seqs, 19947517 residues

Total number of hits satisfying chosen parameters: 133702

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA New:
1: /cgn2_6/ptodaca/1/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodaca/1/pubpaa/US06_NEW_PUB pep:*
3: /cgn2_6/ptodaca/1/pubpaa/US07_NEW_PUB pep:*
4: /cgn2_6/ptodaca/1/pubpaa/PCR_NEW_PUB pep:*
5: /cgn2_6/ptodaca/1/pubpaa/US09_NEW_PUB pep:*
6: /cgn2_6/ptodaca/1/pubpaa/US10_NEW_PUB pep:*
7: /cgn2_6/ptodaca/1/pubpaa/US11_NEW_PUB pep:*
8: /cgn2_6/ptodaca/1/pubpaa/US60_NEW_PUB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1482.5	64.7	406	7	US-11-072-512-3928
2	90	3.9	800	7	US-11-024-959-511
3	90	3.9	984	6	US-10-995-561-629
4	89	3.9	452	7	US-11-087-099-7273
5	88	3.9	800	7	US-11-024-959-471
6	88	3.8	1345	7	US-11-052-554-282
7	87.5	3.8	1121	7	US-11-024-959-459
8	87	3.8	320	7	US-11-087-099-7278
9	87	3.8	2105	7	US-11-052-554-173
10	86.5	3.8	266	7	US-11-072-512-3891
11	86	3.8	759	7	US-11-149-003-22
12	86	3.8	1057	7	US-11-149-003-6
13	86	3.8	1192	7	US-11-149-003-18
14	86	3.8	1207	7	US-11-149-003-20
15	86	3.8	1251	7	US-11-149-003-16
16	86	3.8	1342	7	US-11-149-003-24
17	86	3.8	1477	7	US-11-149-003-8
18	86	3.8	1512	7	US-11-149-003-10
19	86	3.8	1535	7	US-11-149-003-14
20	86	3.8	1570	7	US-11-149-003-12
21	86	3.8	1593	7	US-11-149-003-4
22	86	3.8	1628	7	US-11-149-003-2
23	86	3.8	2176	7	US-11-193-561-25
24	86	3.8	2176	7	US-11-193-771-25
25	86	3.8	2176	7	US-11-193-789-25

26	86	3.8	2176	7	US-11-193-806-25	Sequence 25, Appl
27	86	3.8	2176	7	US-11-193-857-25	Sequence 25, Appl
28	86	3.8	2217	7	US-11-193-561-38	Sequence 38, Appl
29	86	3.8	2217	7	US-11-193-771-38	Sequence 38, Appl
30	86	3.8	2217	7	US-11-193-789-38	Sequence 38, Appl
31	86	3.8	2217	7	US-11-193-806-38	Sequence 38, Appl
32	86	3.8	2217	7	US-11-193-857-38	Sequence 38, Appl
33	86	3.8	2223	7	US-11-193-561-2	Sequence 2, Appl
34	86	3.8	2223	7	US-11-193-771-2	Sequence 2, Appl
35	86	3.8	2223	7	US-11-193-789-2	Sequence 2, Appl
36	86	3.8	2223	7	US-11-193-806-2	Sequence 2, Appl
37	86	3.8	2223	7	US-11-193-857-2	Sequence 2, Appl
38	86	3.8	2296	6	US-10-995-561-633	Sequence 633, Appl
39	86	3.8	2296	7	US-11-193-561-23	Sequence 23, Appl
40	86	3.8	2296	7	US-11-193-771-23	Sequence 23, Appl
41	86	3.8	2296	7	US-11-193-789-23	Sequence 23, Appl
42	86	3.8	2296	7	US-11-193-806-23	Sequence 23, Appl
43	86	3.8	2296	7	US-11-193-857-23	Sequence 23, Appl
44	86	3.8	2330	7	US-11-193-561-21	Sequence 21, Appl
45	86	3.8	2330	7	US-11-193-771-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-11-072-512-3928
Sequence 3928, Application US/11072512
Publication No. US2006002945A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGITAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKU
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3928
LENGTH: 406
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-3928
Query Match 64.7%; Score 1482.5; DB 7; Length 406;
Best Local Similarity 65.6%; Pred. No. 4.8e-130;
Matches 275; Conservative 36; Mismatches 65; Indels 43; Gaps 4;
Cy 3 SPCGEHCAHPKPKYKGLVLYGYNALPNDGRGRKSRFALYKPKKANGVSPVTHVI 62
Db 30 SPG-EDAGGEHPKPKYKGL-----I 48
Cy 63 STPOASKALSKCKQSHISTYLSRNQTVVVEYTHDKDTDFQVGKSTESPIDFVYDTISG 122

Db	49	STPLVSKALSNNGQHSISITLRSKSHVIEYHTSDSTDMQJQRSTENMIDPVYDTSPG	108
Qy	123	SQNTDEAOITQSTISRACRIVCDNEPYTARIIPAAGPSSKNIPTJGEKAKKMPDGHM	1622
Db	109	G-GAAEGSPASOSTISIRYACRIILCDRRPPYARIYAAGFADSSNIPLGEBAKARTPDGLM	167
Qy	163	DGLTTNGVLVMMPRGPFTEESOPGWREISVCGDYVTLTETBSA00RGKIVES2TVLOD	242
Db	168	DGLTTNGVLVMMPRGPFTEESOPGWREISVCGDYVTLTETBSA00RGKIVES2TVLOD	227
Qy	243	GSLIDLCGATLLMRTADGLFHPTQKHIEALROEINAAAPQCCVGLNTLAPSPINKEVY	3022
Db	228	GSLIDLCGATLLMRTAPAGLLRAPLTLQLEAQREANAAARQCVCVGLSTLAPSPINKEVY	287
Qy	303	BEK-OPMAYLSCGHVGHYHNMGRSTEANERECPMCRITVGPVYPLMLGCEAGPYVDAGP	361
Db	288	PDQOQPMVYVVRGCHVGHYHNMGRGRERBCPOERBCPCRLTGPVYPLMLGCEAGLCLDPGP	347
Qy	362	FTHAFTPCGHVSCSEKSAKWSQIPLPHGTHAFHAPACDFCATOLVGEONCIKLLPOSGID	420
Db	348	PSHAAPFCGHVSCSEKARTYMAQCPPLPHGTHAFHAPACDFCATOLVGEONCIKLLPOSGID	406

```

RESULT 2
US-11-024-959-511
; Sequence 511, Application US/11024959
; Publication No. US20060010516A1
GENERAL INFORMATION:
APPLICANT: FORSTER, RICHARD L.
APPLICANT: CONNETT, MARIE B.
APPLICANT: EMERSON, SARAH JANE
APPLICANT: GRIGOR, MURRAY ROBERT
APPLICANT: HIGGINS, COLLEEN M.
APPLICANT: LUND, STEVEN TROY
APPLICANT: MAGOSIN, ANDREAS
APPLICANT: KODRZYCKI, BOB
TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
FILE REFERENCE: 044463-0360
CURRENT APPLICATION NUMBER: US/11/024,959
CURRENT FILING DATE: 2004-12-30
PRIOR APPLICATION NUMBER: 60/533,036
PRIOR FILING DATE: 2003-12-30
NUMBER OF SEQ ID NOS: 782
SOFTWARE: PatentIn version 3.3
SEQ ID NO 511
;
; LENGTH: 800
;
; TYPE: PRT
;
; ORGANISM: Pinus radiata
;
US-11-024-959-511

```

	Query Match	3.9%	Score 90;	DB 7;	Length 800;
	Best Local Similarity	26.4%	Pred. No. 3,6;		
	Matches	46;	Conservative	9;	Mismatches 59; Indels 60; Gaps 6.
Oy	47	KRPKANGKPSTVHISTPO-ASKAISCKG-----QHSIS-----YT	82		
Dd	351	KEPKTTSLTITBGLSSTPQRAGIAFSKKNLPASSGPSYSTPKKNSTRHQPTTNQT	410		
Oy	83	LSRNQTVVV-----EYTHDKOTMPOVGASTESPIDFVTDTTSGSONDE---	128		
Dd	411	LSRPDIIVEIVPRNSRLPETTSDDAKCMNNNGRAVPVSTVSRTSKSDVIKSGSNRDESDKI	470		
Oy	129	-----AQTOSTISR-----ACRIODNREPYTAIRIIFAFG	160		
Dd	471	DSINOKMTGNDKTDLNTARAEQHVSHSRDINTNTISSVVCDOGAPARNIGAAKF	524		

RESULT 3
US-10-995-561-629
; Sequence 629, Application US/10995561
; Publication No. US2005027205A1
; GENERAL INFORMATION:

```

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE. METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 629
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-629

```

[illegible]

```

RESULT 4
US-11-087-099-7273
; Sequence 7273, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B BP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7273
; LENGTH: 452
; TYPE: PRY
; ORGANISM: Nostoc punctiforme
US-11-087-099-7273

```

```

Query Match          3.9%, Score 89; DB 7; Length 457;
Best Local Similarity 22.1%, Pred. No. 2;
Matches 79; Conservative 32; Mismatches 149; Indels 98; Gaps 14.

QY 48 RPKANGKVPSTVHYISTPQASKAISCKGQHSISTLSLN----- 86
DB 51 RKRAVLTGPSNRRVLVEQGAENMSRGVYLSSTFGNNILLQGEHRILTRRLMYPAAH 110

```

```

QY 87 -----QTVVVEYTHD-KDTMFOVGRSTESPIDFVVTDTTSGSONTDEAOITQS 134
DB 111 GKAIATVPDITQIVQDFLDMKMERGISLSSFRQITLMIATRLPLSGQNKSEVE----- 166
QY 135 TISFACRIYCDRNEPYTARIIPAAGPSSKNIFLGEKAARKNP-DGHMDGLTTNGVLM 193
DB 167 -----QTSQMFQTLDDSSMAIF-----KMNVPFTLYGRGQVARGGLVA 204
QY 194 HPRGCTEESQPGVMEISVCGVYTLRTFSRNOQRKQVSESTNVLDGSLDLCG--- 250
DB 205 FLEAIAQRIBQGNLESK---DVLGLLLAVIDEDGKLT--SEQVINEALLLPAGHET 259
QY 251 -ATLWRADGLHTPTOKHIEALROBINARPOCPVGLNTL-APPSI-NRKEVEEKOP 307
DB 260 TABLLTVMVIFELGNHPEMR--ERLRQQLAVGNNPLSLHLQFQPOLTVLKEAEKLYP 317
QY 308 MAYLSCGHVGHYNNHGRSDTEANERCEPMCRVTGVPYPLMLCCEAGFYVDAGPP-TH 364
DB 318 PVYA-----YNRGVLKDIERYGVRIIP-----AGMFTTISPMPLTH 351

```

RESULT 5

```

US-11-024-959-471
; Sequence 471, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: Patentin version 3.3
; SEQ ID NO. 471
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-471

```

Query Match 3.9%; Score 89; DB 7; Length 800;

Best Local Similarity 26.4%; Pred. No. 4.5; Matches 46; Conservative 9; Mismatches 59; Indels 60; Gaps 6;

```

QY 47 KRPKANGVKSPTVHVISTPQ-ASKAISCKG-----CHSIS-----YT 82
DB 351 KEKRTTSLTTBGLSTPQAGAFBSKULPASSGPPSVSTPKKNSTSRVQPTTNFOT 410
QY 83 LSNQVTVV-----BYTHDKDTMFQVGRSTESPIDFVVTDTTSGSONTDE----- 128
DB 411 LSRPDIIVPVIPRNSNLRPETSDVKKEMNFGSVSTVSTKSTVDIKSGSRDESDKI 470
QY 129 -----AQITGSTISRF-----ACRVCDBRNEPYTKARIPAAGP 160
DB 471 DGINQKMTGNDKTDLNIARAEOHVSSRLDNTNTSSVVCQGNOPARWIGAKAF 524

```

RESULT 6

```

US-11-052-554A-282
; Sequence 282, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL

```

```

; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: Patentin version 3.3
; SEQ ID NO. 282
; LENGTH: 1345
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-282

```

Query Match 3.8%; Score 88; DB 7; Length 1345;

Best Local Similarity 22.6%; Pred. No. 12; Matches 58; Conservative 32; Mismatches 87; Indels 80; Gaps 13;

```

QY 50 KANGVKSPTVHVISTPQASKAISCKQHISVTLSSNOT-----VVVEY 93
DB 90 QANG-----EAHV--TLKSKA-----GHTVTATLSNNNTSDSQPTFVADKTSALVVLQI 139
QY 94 THDKDT-----DMFOVGRSTESPIDFVVTDT-----ISGSONTDEAOITQST 135
DB 140 SKKEITGNGVDSATLTAIVKQPD-NEVNNLPVTFSTASGLTLTPGSEVTNMSGIAQAT 198
QY 136 ISRFACRIYCDRNEPYTARIIPAAGPSSKNIFLGEKAARKNP-----KNP 178
DB 199 LAG-----VAFGEQVYTAISANNAGSDNKTVPHFIGTAAKIIETLPVPDSIIAGTPONS 253
QY 179 DGHM-----DGLTTNGVLMHPRGCTEESQPG---VMEISVCGVYTLRTFSRNO 227
DB 254 SCSVITATVVDNNGFPVKGYVTFSTNAATVAENTNGQAVTNEQKATVYTT--NTRSSI 311
QY 228 ORGLVSESTNVLDGS 244
DB 312 ESGAPDVTYBASLNGS 328

```

RESULT 7

```

US-11-024-959-459
; Sequence 459, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: Patentin version 3.3
; SEQ ID NO. 459
; LENGTH: 1121
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-459

```

Query Match 3.8%; Score 87.5; DB 7; Length 1121;

Best Local Similarity 20.9%; Pred. No. 10; Matches 98; Conservative 53; Mismatches 143; Indels 175; Gaps 27;

```

QY 32 PNDGRRRSRFPALYKRPKANGVKT-----STVHVISTPQASKA 70

```

Db 207 PNGARPTPTASTGLVLPKATPTPIGHPFPSSSTATSGLASWMSNPMVTSPQAPVA 266
 QY 71 ISCKGHSISYTLRSNQTVEYTHDKDTMFQVGRSTESPIDPVYDTIS-----GS 123
 Db 267 V-----GSEVFPVPPNATILKRP-----RTPPSSSVVDIQTRDSQILKRLRPVS 312
 QY 124 QNTDEA-----QITOST--ISRFACRIVCDRNEPYTARIFAAGPDSKNIFL--GEKA 172
 Db 313 QSIDBATYGPPTLRVPMSTDDLPKTLARAL--NEPVP--VTSIDFHPSQOTFLVGTCKN 367
 QY 173 AK--WKNPGRHMDGLTTNGVLVW-----HPRGGTESQPGVWRREISVCGDVYTLRE 222
 Db 368 GEITLWE--VGSREKCATRSFKIMDNANCSNHLPAALVSVSINK-----412
 QY 223 TRSAQORGLVSESTVNL--QDGLIDLCGATLLMTADGTFHTPOKHTEALRQ--EIN 278
 Db 413 -----VLMSPDGLIGIA-----FTKHLVHTYTQGLD-LRQHLIED 448
 QY 279 AARPOCPVGLNTLAFPSINKREVEERKQPMAYLSCG-----VHG--YHMGHRS 326
 Db 449 AH-----VCGVNDLAFSHPNQQLCV-----VTGGDDKMIKWDATVGRKLVNFBGHDA 496
 QY 327 DTEANEREC-----MCRITYGPVYPLMLGCEAGFYVDAGPPTHAFTPCGHVSEKS 377
 Db 497 PVS--VCPHHKENTQIFSTAVDGKIKAMLYDHLGSRVDYDAP-----GHSTCT--543
 QY 378 AKTWSQILPHGTHAFHACPFCAQTQVGE-----QNCITLIRFG 417
 Db 544 -----TMMYSADGTRLFS-----CGTSKSGESFLVENHSGALIKRTYSG 583

RESULT 8

US-11-087-099-7278
 ; Sequence 7278, Application US/11087099
 ; Publication No. US20060041961A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark S. et al.
 ; TITLE OF INVENTION: Genes and Uses for Plant Improvement
 ; FILE REFERENCE: 38-21(53450)B BP
 ; CURRENT APPLICATION NUMBER: US/11/087,099
 ; CURRENT FILING DATE: 2005-03-22
 ; NUMBER OF SEQ ID NOS: 12464
 ; SEQ ID NO 7278
 ; LENGTH: 320
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(320)
 ; OTHER INFORMATION: unsure at all Xaa locations
 US-11-087-099-7278

Query Match 3.8%; Score 87; DB 7; Length 320;

Best Local Similarity 22.1%; Pred. No. 1.9; Matches 89; Conservative 42; Mismatches 135; Indels 136; Gaps 24;

QY 28 NGALPNDGRGRKSRFALYKRPKANGVSPVYHISTPQASK-AISCKGHSISYTLRSN 86
 Db 11 NLRPLRPESEBRPRPLPPTTTIAKPSAGDTIASADLEKLAVLGHGNGGVYKX-RH 69
 QY 87 QTVVEYTHDKDTMFQVGRSTESPIDVDTISGQNTDEAQITOSTISRFA--CRIV 144
 Db 70 KITSATYA-----LKIHSN--ADATRRRAFSETSLIKRATDC---106
 QY 145 CDRNEPYTARIFAAGPDSKNIFLGEKAAKKNPDG-----HMDGLTTNGVLVWHPRG 197
 Db 107 -----PHVVPFHS-----FENSGVAALIMETWGGTLETALA--TGG 143
 QY 198 GPTESQPGVWRREISVCGDVYTLLET--RSAQORGLVSESTVNLQDGLIDLCGATLL 254
 Db 144 TFSERLAKYARV-LEGLAVLHAARNTAHRDIKIPANTLVNSEGV-----KIADGVSKLM 198
 QY 255 WRTADGLFHTPTQGHTEALRQEIWAARQCPVGLNTLAFPSINKREVEERKQPMAYLSCG 314

Db 199 CRT-----LEA-----CNSYVGCAYMSPDPRD-----PEAY--CG 227
 QY 315 HNAGY--HNW-----GHRSDTEANERECRVCRTVGPVPLW--LGCAGFYVD- 358
 Db 228 NYNGFADWISGLTLFELYVGHFPLQAGQR-----PDMATLMCAICFSDP 275
 QY 359 -----ACPTHAFTPCGHVSEK-SAKTWSQILPHGTHAF 393
 Db 276 XSLPRTASRPFHDFVFC-----CLKESGERMTAQLL--THPF 312

RESULT 9

US-11-052-554A-173
 ; Sequence 173, Application US/11052554A
 ; Publication No. US2005028866A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sachdeva, et al.
 ; TITLE OF INVENTION: COMBINATION METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
 ; FILE REFERENCE: 30853/40359A
 ; CURRENT APPLICATION NUMBER: US/11/052,554A
 ; CURRENT FILING DATE: 2005-02-07
 ; PRIOR APPLICATION NUMBER: US 60/589,227
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
 ; PRIOR FILING DATE: 2004-02-06
 ; NUMBER OF SEQ ID NOS: 763
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 173
 ; LENGTH: 2105
 ; TYPE: PRT
 ; ORGANISM: Porphyromonas gingivalis W83
 US-11-052-554A-173

Query Match 3.8%; Score 87; DB 7; Length 2105;

Best Local Similarity 22.9%; Pred. No. 27; Matches 56; Conservative 35; Mismatches 86; Indels 68; Gaps 17;

QY 4 PQGEHCAPKKEPVYKGLVYLGYNALPNG-DGRGRKSRFALYKRPKANGVSPVYH- 61
 Db 714 PGHVAWMSKSTGNAGDFTVFEE--TPNGINKG--GAFFGL--STEANGARQSWALR 767
 QY 62 --ISTPOASKAISCKQH--SISYTLRSNQTVEYTHDKDTMFQVGRSTESPIDPVY 117
 Db 768 RYVDLPAGTKYVAFRHNCGDLATYL-----LDQITMGGS-PFTTYT 813
 QY 118 DTISGQNTDEAQITOSTISRFACRIVCDRNEPYTARI-FAAG-----FD 161
 Db 814 VYRDGFKIKE--GLTETTFEEDG---VATGNHEYCVKVTAGVSPKECVNVTVDPVQFN 868
 QY 162 SKNT--FLGEKAA-KMKNPDGHMDGLTTNGVLVWHPRGFT-----EESQPGVWRREI 211
 Db 869 PVNLGSAVAGQVTKLMDAPNG-----FPN-----PMDGTTLSSEFENGIPASWKT 917
 QY 212 SVCGD 216
 Db 918 DADGD 922

RESULT 10

US-11-072-512-3891
 ; Sequence 3891, Application US/11072512
 ; Publication No. US2006002945A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO

```

APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: MAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHIKO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072,512
PENDING FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3891
LENGTH: 266
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-3891

```

	Query Match	3.8%;	Score 86;	DB 7;	Length 759;	
	Best Local Similarity	20.5%;	Pred. No. 7.9;			
	Matches	64;	Conservative	28;	Mismatches 120;	Indels 100;
					Gaps	17
Qy	166	I F G E K A K A W K N P	-----D G H M D G L	-----	T T N G V L V M H P R G G F	-----T 200
Db	463	V I L D G M A V R L L Q D G A V T V D G H P V A L P F L Q E B L T Y L E R G H T V I L A Q G L Q V L M D G S Q S V				522
Qy	201	E E S Q P G T W	--R E I S V C D V Y T L R T R S A Q Q G K V E S F T N V L Q D G S L I D L C G A T I L M T R T A			258
Db	523	E V S V S G S Y Q G T C G L C A N F N G F A D D I L O G P E G L L P S E A A F G N S	-----		W O V S	570
Qy	259	D G L P	--H T P Q K E I A L R E I N A R P O C P V G I L T L A P F S I N R K E V E E K O P	-----	W A	309
Db	571	E G L M P G R P C S A G R E V D P C P R A G Y A R R E A N N A R C G V L S S P F S R C H A V V P P R E F P A A C Y I D				630
Qy	310	Y L S C G H V G H Y N M G H R S D T	--E A N E B E C M C R T G V P Y P I A M	-----	L G C	E A G E Y V 357
Db	631	L C A G S P	-----G S S A D A C L C L D L A E V A S H C R Q A G	--T P M W R G E T L C V V C S P L E R G E V F		682
Qy	358	D A C P P	-----T A F T P C G H V C E K S A K Y S Q I P L P R G T A F H A	-----	A	396
Db	683	D E C E P P C P R T C F N G H I P L G E L A A H C V A R P C V R G C	-----		O C P A G L V E R H A R C I P P B A	733
Qy	397	C P F C A T Q L V G E Q				408
Db	734	C P	--Q V L I T G D Q			743

[illegible]

```

RESULT 13
US-11-149-003-18
; Sequence 18, Application US/11149003
; Publication No. US20060014277A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Novel Human Kiflin-like Proteins and Polynucleotides Encoding the
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0360-USA
; CURRENT APPLICATION NUMBER: US/11/149, 003
; CURRENT FILING DATE: 2005-06-09
; PRIOR APPLICATION NUMBER: US/10/189,971
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/302,949
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/315,634
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1192
; TYPE: PRT
; ORGANISM: homo sapiens
; US-11-149-003-18

```

Query Match	3.8%	Score 86;	DB 7;	Length 1192;
Best Local Similarity	20.5%	Pred. No. 15;		
Matches	64;	Conservative	28;	Mismatches 120;
			Indels 100;	Gaps 17;

QY	166	IFLGEKAKWKNP-----DGHMDGL-----TTNGVLVWNRGQF-----T	200
Db	896	VLLDGMAYRLLQDGAVTVDGHPVALPFLQEPRLLYELRGHTYILHAQGLQVLMQDSQSV	955
QY	201	BESQGVW--REISVGDVYTLIRETRSAQQRKGLVESETNVVLQDSSLIDLGGATTLNRTA	258
Db	956	EVSVPGSYQGRTCGICGNFNGPAPQODLQGPBGLLPSBAAFNS-----WQVS	1003
QY	259	DGLF---HTPQKHLEALRQELINARPCQPVGLNTLAPRINRKEVNEBKQ-----WA	309
Db	1004	FGLMPGRCSAGREVPDPCRAAGYRARREANARCGVLKSSPFSRCHAVVPPEFPFACAYD	1063
QY	310	YLSCGVHGVHNMGRSDT--EANERBCPMCRITGPPVPLW-----LGC--EAGEYV	357
Db	1064	LCAGCP-----GSSADACLDALBAASHRQNG-VTPTRGRGTLTLCVCGCPLEGRIVF	1115
QY	358	D-ACGP-----THAFTPCGHVCSKSAKYSQIDLPHTHAFAH---A	396
Db	1116	DECGPPCPRTCFNGHIFLGLBLAAHCVRCVPCG-----QCPAGVHEHAHCIPEBA	1166
QY	397	CPFCATQLVGEQ	408
Db	1167	CP--QVLLTGDQ	1176

RESULT 14
US-11-149-003-20
; Sequence 20, Application US/11149003
; Publication No. US20060014277A1

```

GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: Novel Human Kiehn-like Proteins and Polynucleotides Encoding the
FILE OF INVENTION: Same
FILE REFERENCE: LEX-0360-USA
CURRENT APPLICATION NUMBER: US/11/149,003
CURRENT FILING DATE: 2005-06-09
PRIORITY APPLICATION NUMBER: US/10/189,971
PRIORITY FILING DATE: 2002-07-03
PRIORITY APPLICATION NUMBER: US 60/302,949
PRIORITY FILING DATE: 2001-07-03
PRIORITY APPLICATION NUMBER: US 60/315,634
PRIORITY FILING DATE: 2001-08-29
NUMBER OF SEQ. ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 20
LENGTH: 1207
TYPE: PRT
ORGANISM: homo sapiens
US-11-149-003-20

Query Match          3.8%; Score 86; DB 7; Length 1207;
Best Local Similarity 20.5%; Pred. No. 15;
Matches 64; Conservative 28; Mismatches 100; Gaps 17;

```

	Query March	3.8%; Score 86; DB 7; Length 1207;
	Best Local Similarity	20.5%; Pred. No. 15;
	Matches 66; Conservative	26; Mismatches 120; Indels 100; Gaps 17;
QY	166 IFLEKKAAXKNP-----DGHMDGL-----TTNGVLVNHPRGCF-----T	200
Db	911 VILGDMAVRLLDDGAIVTDGHPVALPTLGEPLLYVELRGHTVILHAQPGQLVMDSQS	970
QY	201 EESOPGW--REISVGSDVYTLLETBSAQQRKGLVBSETVNLDDSLIDLCSATLLMRTA	258
Db	971 EVSPSGSYOQRTCLCGNFNGFADDDIQGPBGILLPSBEAAFGNS-----MQVS	1018
QY	259 DGLP----HTPQKIIEALROEINARPCQCVGLNTLLAPSIINKEVVEBKQ-----WA	309
Db	1019 EGMLPPGPCSAGRVDPCCRAAGTRARRANARCCVLASSPPSRCHAVVPBPFPACAYTD	1078
QY	310 YLSGGHYGVHNMGRSDT---EAMRECPMCRTVGVPVLM-----LGC-EAGEYV	357
Db	1079 LCAQGF-----GSSADACLCDLLEVAASHCRAG-VTPWRKPRTLCAVSGPLERGVF	1130
QY	358 D-AGRP-----THAFTPCGHVCSEKSAKYWSQIPLPHGTAFPA---A	396
Db	1131 DECSPRCPTCFNQHIFLGEIAAACVRRPVGCF-----CCPAGIVEAHACPPEA	1181
QY	397 CPFCATQLVGEQ	408
Db	1182 CP-QVLLTGDO	1191

```

RESULT 15
US-11-149-003-16
: Sequence 16: Application US/11149003
: Publication No. US20060014277A1
: GENERAL INFORMATION:
: APPLICANT: Walke, D. Wade
: APPLICANT: Scoville, John
: APPLICANT: Turner, C. Alexander Jr.
: TITLE OF INVENTION: Novel Human Kiehn-1-like Proteins and Polynucleotides Encoding the
: TITLE OF INVENTION: Same
: FILE REFERENCE: LEX-0360-USA
: CURRENT APPLICATION NUMBER: US/11/149,003
: CURRENT FILING DATE: 2005-06-09
: PRIOR APPLICATION NUMBER: US/10/189,971
: PRIOR FILING DATE: 2002-07-03
: PRIOR APPLICATION NUMBER: US 60/302,949
: PRIOR FILING DATE: 2001-07-03
: PRIOR APPLICATION NUMBER: US 60/315,634
: PRIOR FILING DATE: 2001-08-29
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 4.0

```


SEQ ID NO 16
LENGTH: 1251
TYPE: PRT
ORGANISM: homo sapiens
US-11-149-003-16

Query Match 3.8%; Score 86; DB 7; Length 1251;
Best Local Similarity 20.5%; Pred. No. 16;
Matches 64; Conservative 28; Mismatches 120; Indels 100; Gaps 17;

QY	166	IFLGEKAKAKWNP	-----DGHMDGL-----	TTNGVLVMEPRGGF-----	T	200
DB	955	VLLGDMAVRLIQDGAVTVDGHPVALPFLQBPILVYELRGHTVILHAQPGLOVMDGQSOV				1014
QY	201	EEGQPGTW--REISVCGDVYTLRETRSAQQRKLYESETNVLQDGLIDLGCATILMRTA				258
DB	1015	EVSVPGSYGRTCGICGNGFNQPAQDDLOGPBGILLPSEAFQNS				1062
QY	259	DGLF--HTPTQKHLEALQEIINARPCQPVGLNTLAFPSINRKVEVEKOP			WA	309
DB	1063	EGLMFGRPCSGAGREVDPCRAAGYRAAREANARCGVLKSSPFSRCHAVVPEPFPAACYID				1122
QY	310	YLSGCVHVGYNWCHRSPT--EANRECEPMCRVTGVPYPLM			IGC--EAGFYV	357
DB	1123	LCACGP-----GSSADACLDALAEAVASHCRQAG--VTPWRGPYLCVVCPIERGFVF				1174
QY	358	D-AGPP-----	THAFTPCGHVCEKSAKYSQITPLPHGTHAFHA		A	396
DB	1175	DECGPPCPRTCFNOHIPLAGELAAHCVRCVPGC-----	QCPAGLVEHEAHCIPEEA			1225
QY	397	CPFCATOLVGEQ				408
DB	1226	CP--QVLLITDQ				1235

Search completed: March 1, 2006, 21:04:59
Job time : 21 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 21:01:01 ; Search time 167 Seconds
(without alignments)
1050.828 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MESPGEHRCAPKPEPVYKXG.....ATQLVGEQNCIKILFQGPID 420

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications AA Main:
1: /cgn2_6/ptodaca/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodaca/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodaca/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodaca/1/pubpaa/US10_PUBCOMB.pep:*
5: /cgn2_6/ptodaca/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodaca/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2290	100.0	420	US-10-041-030-4	Sequence 4, Appl
2	2290	100.0	420	US-10-197-666A-136	Sequence 136, App
3	2283	99.7	420	US-09-843-905A-8	Sequence 8, Appl
4	2283	99.7	420	US-10-317-250-8	Sequence 8, Appl
5	2283	99.7	420	US-10-258-703-8	Sequence 8, Appl
6	2175.5	95.0	419	US-09-843-905A-6	Sequence 6, Appl
7	2175.5	95.0	419	US-10-317-250-6	Sequence 6, Appl
8	2175.5	95.0	419	US-10-258-703-6	Sequence 6, Appl
9	2111.5	92.2	419	US-10-197-666A-134	Sequence 134, App
10	1917	83.7	418	US-09-843-905A-4	Sequence 4, Appl
11	1917	83.7	418	US-10-317-250-4	Sequence 4, Appl
12	1917	83.7	418	US-10-443-108-10	Sequence 10, Appl
13	1917	83.7	418	US-10-258-703-4	Sequence 4, Appl
14	1913	83.5	418	US-10-041-030-2	Sequence 2, Appl
15	1912	83.5	418	US-09-843-905A-2	Sequence 2, Appl
16	1912	83.5	418	US-10-317-250-2	Sequence 2, Appl
17	1912	83.5	418	US-10-258-703-2	Sequence 2, Appl
18	1642	71.7	445	US-09-843-905A-12	Sequence 12, Appl
19	1642	71.7	445	US-10-317-250-12	Sequence 12, Appl
20	1642	71.7	445	US-10-258-703-12	Sequence 12, Appl
21	1482.5	64.7	406	US-10-104-047-3928	Sequence 3928, App
22	1302	56.9	424	US-09-843-905A-13	Sequence 13, Appl
23	1302	56.9	424	US-10-317-250-13	Sequence 13, Appl
24	1302	56.9	424	US-10-258-703-13	Sequence 13, Appl
25	1302	56.9	424	US-11-097-143-10851	Sequence 10851, A
26	1235	53.9	455	US-09-843-905A-14	Sequence 14, Appl
27	1235	53.9	455	US-10-317-250-14	Sequence 14, Appl

28	1235	53.9	455	4	US-10-258-703-14	Sequence 14, Appl
29	1077	47.0	241	5	US-10-450-763-34853	Sequence 34853, A
30	974	42.5	171	4	US-10-029-386-32174	Sequence 32174, A
31	903	39.4	458	3	US-09-843-905A-15	Sequence 15, Appl
32	903	39.4	458	4	US-10-317-250-15	Sequence 15, Appl
33	903	39.4	458	4	US-10-258-703-15	Sequence 15, Appl
34	775	33.8	155	3	US-09-867-550-1948	Sequence 1948, Ap
35	678	29.6	125	3	US-09-867-550-1720	Sequence 1720, Ap
36	343	15.0	528	5	US-10-450-763-34852	Sequence 34852, A
37	154	6.7	26	4	US-10-041-030-38	Sequence 38, Appl
38	113	4.9	818	5	US-10-732-923-14186	Sequence 14186, A
39	113	4.9	818	5	US-10-732-923-14187	Sequence 14187, A
40	109	4.8	1200	5	US-10-450-763-36074	Sequence 36074, A
41	109	4.8	1379	5	US-10-450-763-36078	Sequence 36075, A
42	104	4.5	20	4	US-10-041-030-26	Sequence 26, Appl
43	100	4.4	19	4	US-10-041-030-31	Sequence 31, Appl
44	100	4.4	1114	5	US-10-732-923-14185	Sequence 14185, A
45	100	4.4	2042	4	US-10-192-584-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-10-041-030-4
Sequence 4, Application US/10041030
Publication NO. US20020150934A1
GENERAL INFORMATION:
APPLICANT: Powers, Scott
APPLICANT: Mr. David
APPLICANT: Xiang, Phil
APPLICANT: Peng, Yue
APPLICANT: Tularik Inc.
TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
FILE REFERENCE: 018781-006810US
CURRENT APPLICATION NUMBER: US/10/041, 030
CURRENT FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: US 60/259,502
PRIOR FILING DATE: 2001-01-02
NUMBER OF SEQ ID NOS: 42
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 420
TYPE: PRT
ORGANISM: Homo sapiens
OTHER INFORMATION: human pellino 2
US-10-041-030-4
Query Match 100.0%; Score 2290; DB 4; Length 420;
Best Local Similarity 100.0%; Pred. No. 1.6e-216;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MESPGEHRCAPKPEPVYKXGELVLYGNGALPMGDRGRKSRALYKRPANVKSSTVH 60
1 MESPGEHRCAPKPEPVYKXGELVLYGNGALPMGDRGRKSRALYKRPANVKSSTVH 60
VITPQSKAISCKGHSISYTLRSNQTQVVEYTHKQDTMDFVGRSTESPIFVVDTI 120
61 VITPQSKAISCKGHSISYTLRSNQTQVVEYTHKQDTMDFVGRSTESPIFVVDTI 120
61 VITPQSKAISCKGHSISYTLRSNQTQVVEYTHKQDTMDFVGRSTESPIFVVDTI 120
121 SGGQNTDEAQITQSTISRFACRIVCDRNEBYTARIPAGFSSKNIFLGKAAKXNPDG 180
121 SGGQNTDEAQITQSTISRFACRIVCDRNEBYTARIPAGFSSKNIFLGKAAKXNPDG 180
121 SGGQNTDEAQITQSTISRFACRIVCDRNEBYTARIPAGFSSKNIFLGKAAKXNPDG 180
181 HMGGLTNGVLVHMPRGFTBESQPGVWREISVCGVYTLRFRSAQQRKLVESNTVL 240
181 HMGGLTNGVLVHMPRGFTBESQPGVWREISVCGVYTLRFRSAQQRKLVESNTVL 240
241 QDPSLIDLCATLMTPTADGLFTTPQKITEALROEINARPOCPGINTLAPPSINRKE 300
241 QDPSLIDLCATLMTPTADGLFTTPQKITEALROEINARPOCPGINTLAPPSINRKE 300

QY 301 VVEKOPWALISCGHVGYNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 301 VVEKOPWALISCGHVGYNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 420
Db 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 420

RESULT 2

US-10-197-666A-136
; Sequence 136, Application US/10197666A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAMI KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: Elki phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 136
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-197-666A-136

Query Match 100.0%; Score 2290; DB 4; Length 420;
Best Local Similarity 100.0%; Pred. No. 1.6e-216;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSPGQEHCAPKKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTYH 60
Db 1 MFSPGQEHCAPKKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTYH 60
QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDIT 120
Db 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDIT 120
QY 121 SSGQNTDEAOITOSTISRFAICRVCDRNEPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
Db 121 SSGQNTDEAOITOSTISRFAICRVCDRNEPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
QY 181 HMDGLTTNGVLVNHPRGFTFEESOPGWREISVCGDVYTLRETRSAOQRKLVESSETNVL 240
Db 181 HMDGLTTNGVLVNHPRGFTFEESOPGWREISVCGDVYTLRETRSAOQRKLVESSETNVL 240
QY 241 QDSGLIDLCGATILMRTADGLFHTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 241 QDSGLIDLCGATILMRTADGLFHTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 300
QY 301 VVEKOPWALISCGHVGYNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 301 VVEKOPWALISCGHVGYNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 420
Db 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 420

RESULT 3

US-09-843-905A-8
; Sequence 8, Application US/09843905A
; Patent No. US2002016683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-843-905A-8

Query Match 99.7%; Score 2283; DB 3; Length 420;
Best Local Similarity 99.8%; Pred. No. 7.7e-216;
Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFSPGQEHCAPKKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTYH 60
Db 1 MFSPGQEHCAPKKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTYH 60
QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDIT 120
Db 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDIT 120
QY 121 SSGQNTDEAOITOSTISRFAICRVCDRNEPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
Db 121 SSGQNTDEAOITOSTISRFAICRVCDRNEPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
QY 181 HMDGLTTNGVLVNHPRGFTFEESOPGWREISVCGDVYTLRETRSAOQRKLVESSETNVL 240
Db 181 HMDGLTTNGVLVNHPRGFTFEESOPGWREISVCGDVYTLRETRSAOQRKLVESSETNVL 240
QY 241 QDSGLIDLCGATILMRTADGLFHTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 241 QDSGLIDLCGATILMRTADGLFHTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 300
QY 301 VVEKOPWALISCGHVGYNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 301 VVEKOPWALISCGHVGYNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 420
Db 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 420

RESULT 4

US-10-317-250-8
; Sequence 8, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-317-250-8

Query Match 99.7%; Score 2283; DB 4; Length 420;
 Best Local Similarity 99.8%; Pred. No. 7.7e-216;
 Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFSPGQEBHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPANGVKSTVH 60
 DB 1 MFSPGQEBHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPANGVKSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLSSRNOTVVEYTHDKOTDMFOVGRSTESPIDFVVDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLSSRNOTVVEYTHDKOTDMFOVGRSTESPIDFVVDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIIPAAGDSSKNITFLGKAKAKMPDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIIPAAGDSSKNITFLGKAKAKMPDG 180
 QY 181 HMDGLTTNGVLVNHPRGFTESQPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240
 DB 181 HMDGLTTNGVLVNHPRGFTESQPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240
 QY 241 QDQSLIDLCGATLLMTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDQSLIDLCGATLLMTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTVGYVPLMGCBAGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTVGYVPLMGCBAGFYVDAG 360
 QY 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFAACPCATQLVGBONCIKLIPOGPID 420
 DB 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFAACPCATQLVGBONCIKLIPOGPID 420

RESULT 5

US-10-258-703-8
 ; Sequence 8, Application US/10258703
 ; Publication No. US20040034199A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-US
 ; CURRENT APPLICATION NUMBER: US/10/258, 703
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 420
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-258-703-8

Query Match 99.7%; Score 2283; DB 4; Length 420;
 Best Local Similarity 99.8%; Pred. No. 7.7e-216;
 Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFSPGQEBHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPANGVKSTVH 60
 DB 1 MFSPGQEBHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPANGVKSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLSSRNOTVVEYTHDKOTDMFOVGRSTESPIDFVVDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLSSRNOTVVEYTHDKOTDMFOVGRSTESPIDFVVDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIIPAAGDSSKNITFLGKAKAKMPDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIIPAAGDSSKNITFLGKAKAKMPDG 180
 QY 181 HMDGLTTNGVLVNHPRGFTESQPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240
 DB 181 HMDGLTTNGVLVNHPRGFTESQPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240

QY 241 QDQSLIDLCGATLLMTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDQSLIDLCGATLLMTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTVGYVPLMGCBAGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTVGYVPLMGCBAGFYVDAG 360
 QY 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFAACPCATQLVGBONCIKLIPOGPID 420
 DB 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFAACPCATQLVGBONCIKLIPOGPID 420

RESULT 6

US-09-843-905A-6
 ; Sequence 6, Application US/09843905A
 ; Patent No. US20020168683A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-A
 ; CURRENT APPLICATION NUMBER: US/09/843, 905A
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,198
 ; PRIOR FILING DATE: 2000-04-28
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 419
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-843-905A-6

Query Match 95.0%; Score 2175.5; DB 3; Length 419;
 Best Local Similarity 95.0%; Pred. No. 3.1e-205;
 Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MFSPGQEBHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPANGVKSTVH 60
 DB 1 MFSPGQEBHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPANGVKSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLSSRNOTVVEYTHDKOTDMFOVGRSTESPIDFVVDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLSSRNOTVVEYTHDKOTDMFOVGRSTESPIDFVVDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIIPAAGDSSKNITFLGKAKAKMPDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIIPAAGDSSKNITFLGKAKAKMPDG 180
 QY 181 HMDGLTTNGVLVNHPRGFTESQPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240
 DB 181 HMDGLTTNGVLVNHPRGFTESQPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240
 QY 241 QDQSLIDLCGATLLMTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDQSLIDLCGATLLMTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTVGYVPLMGCBAGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTVGYVPLMGCBAGFYVDAG 360
 QY 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFAACPCATQLVGBONCIKLIPOGPID 420
 DB 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFAACPCATQLVGBONCIKLIPOGPID 420

RESULT 7

US-10-317-250-6
 ; Sequence 6, Application US/10317250
 ; Publication No. US20030165945A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.

; APPLICANT: Cosman, David J.
 ; APPLICANT: Li, Xiaoxia
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-B
 ; CURRENT APPLICATION NUMBER: US/10/317,250
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 419
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-10-317-250-6

Query Match 95.0%; Score 2175.5; DB 4; Length 419;
 Best Local Similarity 95.0%; Pred. No. 3.1e-205;
 Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFALYKRPKANGVKPSTVH 60
 DB 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFALYKRTYASGVKPSSTIH 60
 QY 61 VISTPQASKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDTI 120
 QY 121 SGSONTDEAQTOSTISRFACRIVCDBNEPYTARIPAAGPSKNIPLGKAAKWKMPDG 180
 DB 121 SGSONTDEAQTOSTISRFACRIVCDBNEPYTARIPAAGPSKNIPLGKAAKWKMPDG 179
 QY 181 HMDGLTTNGVLVNHPRGCFTEESQPGWREISVCGDVYTLRETRSAOQRKLVESSTNVL 240
 DB 180 HMDGLTTNGVLVNHPRGCFTEESQPGWREISVCGDVYTLRETRSAOQRKLVESSTNVL 239
 QY 241 QDGSLLIDCGATTLMTADGLFHTPTQKHIEALROEINARPCPGVLTALAPSINRKE 300
 DB 240 QDGSLLIDCGATTLMTADGLFHTPTQKHIEALROEINARPCPGVLTALAPSINRKE 299
 QY 301 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTGVPVPLMGCEAGFYVDAG 360
 DB 300 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTGVPVPLMGCEAGFYVDAG 359
 QY 361 PPTHAFTPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCKIKLIFGSPVD 420
 DB 360 PPTHAFTPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCKIKLIFGSPVD 419

RESULT 8
 US-10-258-703-6
 ; Sequence 6, Application US/10258703
 ; Publication No. US20040034199A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-US
 ; CURRENT APPLICATION NUMBER: US/10/258,703
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 419
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-10-258-703-6

Query Match 95.0%; Score 2175.5; DB 4; Length 419;
 Best Local Similarity 95.0%; Pred. No. 3.1e-205;
 Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFALYKRPKANGVKPSTVH 60
 DB 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFALYKRTYASGVKPSSTIH 60

QY 61 VISTPQASKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDTI 120
 QY 121 SGSONTDEAQTOSTISRFACRIVCDBNEPYTARIPAAGPSKNIPLGKAAKWKMPDG 180
 DB 121 SGSONTDEAQTOSTISRFACRIVCDBNEPYTARIPAAGPSKNIPLGKAAKWKMPDG 179
 QY 181 HMDGLTTNGVLVNHPRGCFTEESQPGWREISVCGDVYTLRETRSAOQRKLVESSTNVL 240
 DB 180 HMDGLTTNGVLVNHPRGCFTEESQPGWREISVCGDVYTLRETRSAOQRKLVESSTNVL 239
 QY 241 QDGSLLIDCGATTLMTADGLFHTPTQKHIEALROEINARPCPGVLTALAPSINRKE 300
 DB 240 QDGSLLIDCGATTLMTADGLFHTPTQKHIEALROEINARPCPGVLTALAPSINRKE 299
 QY 301 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTGVPVPLMGCEAGFYVDAG 360
 DB 300 VVEKQPMAYLSCGHVGHYHNMGRSDTEANERECPCRTGVPVPLMGCEAGFYVDAG 359
 QY 361 PPTHAFTPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCKIKLIFGSPVD 420
 DB 360 PPTHAFTPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCKIKLIFGSPVD 419

RESULT 9
 US-10-197-666A-134
 ; Sequence 134, Application US/10197666A
 ; Publication No. US20030092037A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ASAH KASBI KABUSIKI KAISYA
 ; TITLE OF INVENTION: Elki phosphorylation related gene
 ; FILE REFERENCE: PH-1548US
 ; CURRENT APPLICATION NUMBER: US/10/197,666A
 ; CURRENT FILING DATE: 2002-11-18
 ; PRIOR APPLICATION NUMBER: JP 2001-218204
 ; PRIOR FILING DATE: 2001-07-18
 ; PRIOR APPLICATION NUMBER: JP 2001-263450
 ; PRIOR FILING DATE: 2001-08-31
 ; PRIOR APPLICATION NUMBER: JP 2002-012176
 ; PRIOR FILING DATE: 2002-01-21
 ; PRIOR APPLICATION NUMBER: US 60/305,884
 ; PRIOR FILING DATE: 2001-07-18
 ; PRIOR APPLICATION NUMBER: US 60/316,304
 ; PRIOR FILING DATE: 2001-09-04
 ; PRIOR APPLICATION NUMBER: US 60/350,027
 ; PRIOR FILING DATE: 2002-01-23
 ; NUMBER OF SEQ ID NOS: 156
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 134
 ; LENGTH: 419
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; NAME/KEY: UNSURE
 ; LOCATION: 357..362
 ; OTHER INFORMATION: unknown
 ; US-10-197-666A-134

Query Match 92.2%; Score 2111.5; DB 4; Length 419;
 Best Local Similarity 92.4%; Pred. No. 6.3e-199;
 Matches 388; Conservative 12; Mismatches 19; Indels 1; Gaps 1;

QY 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFALYKRPKANGVKPSTVH 60
 DB 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFALYKRTYASGVKPSSTIH 60
 QY 61 VISTPQASKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDTI 120
 QY 121 SGSONTDEAQTOSTISRFACRIVCDBNEPYTARIPAAGPSKNIPLGKAAKWKMPDG 180

```

Db 121 SGGNED-AQITGSTISRPACRIYCDNEPTTAIPAAGDSSKNITLGEAAKMPDG 179
Qy 181 HMDGLTNGVLVNHPRGGFTBESOPGWREISVCGDYTTLTRSAOQKGLVESFTNL 240
Db 180 HMDGLTNGVLVNHPRGGFTBESOPGWREISVCGDYTTLTRSAOQKGLVESFTNL 239
Qy 241 QDGLSLIDLCATLMTADGLFHTPTOKHLEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 240 QDGLSLIDLCATLMTADGLFHTPTOKHLEALROEINARPOCPVGLNTLAFPSINRKE 299
Qy 301 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 300 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 359
Qy 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPFCATQLVGBONCIKLIPOGPID 420
Db 360 XXXHVFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPFCATQLVGBONCIKLIPOGPID 419

```

```

RESULT 10
US-09-843-905A-4
; Sequence 4, Application US/09843905A
; Patent No. US20020168683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens.
US-09-843-905A-4

```

Query Match 83.7%; Score 1917; DB 3; Length 418;
 Best Local Similarity 81.4%; Pred. No. 9,4e-180;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

```

Qy 1 MFSPGDEHCAPNKEPVYKGLVYLGYNGALPNDGRGRKSRPALYKRPANGVKSTVH 60
Db 1 MFSPDQENH--PSKAPVYKGLVYLGYNGSLPNDGRGRKSRPALYKRPANGVKSTVH 58
Qy 61 VISTPOASKAISCKGHSISYTLSSRNQTVVEYTHDKDTDMFOVGRSTESPIDFVVTDTI 120
Db 59 IACTPQAARAIKSNKDHISYTLSSRAQTVVEYTHDSNTDMFOIGRSTESPIDFVVTDTV 118
Qy 121 SGGQNTDEAQITGSTISRPACRIYCDNEPTTAIPAAGDSSKNITLGEAAKMPDG 180
Db 119 PGSQNSNDTQSVOSTISRFACRIICERNPPTARIYAAGDSSKNITLGEAAKMPDGS 178
Qy 181 HMDGLTNGVLVNHPRGGFTBESOPGWREISVCGDYTTLTRSAOQKGLVESFTNL 240
Db 179 QMDGLTNGVLVNHPRGGFTBESOPGWREISVCGDYTTLTRSAOQKGLVESFTNL 238
Qy 241 QDGLSLIDLCATLMTADGLFHTPTOKHLEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 239 QDGLSLIDLCATLMTADGLFHTPTOKHLEALROEINARPOCPVGLNTLAFPSINRKE 298
Qy 301 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 299 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 358
Qy 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPFCATQLVGBONCIKLIPOGPID 420
Db 359 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPFCATQLVGBONCIKLIPOGPID 418

```

```

RESULT 11
US-10-317-250-4
; Sequence 4, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; CURRENT FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-317-250-4

```

Query Match 83.7%; Score 1917; DB 4; Length 418;
 Best Local Similarity 81.4%; Pred. No. 9,4e-180;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

```

Qy 1 MFSPGDEHCAPNKEPVYKGLVYLGYNGALPNDGRGRKSRPALYKRPANGVKSTVH 60
Db 1 MFSPDQENH--PSKAPVYKGLVYLGYNGSLPNDGRGRKSRPALYKRPANGVKSTVH 58
Qy 61 VISTPOASKAISCKGHSISYTLSSRNQTVVEYTHDKDTDMFOVGRSTESPIDFVVTDTI 120
Db 59 IACTPQAARAIKSNKDHISYTLSSRAQTVVEYTHDSNTDMFOIGRSTESPIDFVVTDTV 118
Qy 121 SGGQNTDEAQITGSTISRPACRIYCDNEPTTAIPAAGDSSKNITLGEAAKMPDG 180
Db 119 PGSQNSNDTQSVOSTISRFACRIICERNPPTARIYAAGDSSKNITLGEAAKMPDGS 178
Qy 181 HMDGLTNGVLVNHPRGGFTBESOPGWREISVCGDYTTLTRSAOQKGLVESFTNL 240
Db 179 QMDGLTNGVLVNHPRGGFTBESOPGWREISVCGDYTTLTRSAOQKGLVESFTNL 238
Qy 241 QDGLSLIDLCATLMTADGLFHTPTOKHLEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 239 QDGLSLIDLCATLMTADGLFHTPTOKHLEALROEINARPOCPVGLNTLAFPSINRKE 298
Qy 301 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 299 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 358
Qy 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPFCATQLVGBONCIKLIPOGPID 420
Db 359 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHAFHAACPFCATQLVGBONCIKLIPOGPID 418

```

```

RESULT 12
US-10-443-108-10
; Sequence 10, Application US/10443108
; Publication No. US20040005615A1
; GENERAL INFORMATION:
; APPLICANT: LI, JING
; APPLICANT: KU, DAVID
; APPLICANT: YANG, JIANXIN
; TITLE OF INVENTION: AMPLIFICATION AND OVEREXPRESSION OF ONCOGENES
; FILE REFERENCE: 38002-0049
; CURRENT APPLICATION NUMBER: US/10/443,108
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: 60/398,099
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/382,606
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10

```

LENGTH: 418
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-443-108-10

Query Match 83.7%; Score 1917; DB 4; Length 418;
 Best Local Similarity 81.4%; Pred. No. 9,46-180;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGQEHCAPNKEPVYKGLVYLGVNGALPNDGRGRKSRPALYKRPKANGKVPSTVH 60
 Db 1 MFSPDQENH--PSKAPVKYKGLVILVINGSLPNDGRGRKSRPALFKRPKANGKVPSTVH 58
 QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDTDMFQVGRSTESPIDFVYTDI 120
 Db 59 IACTPQAKAISNDQHSISYTLSSRAQTVVVEYTHDSNTDMFQIGRSTESPIDFVYTDV 118
 QY 121 SGSQNTDEAQTOSTISRFACRIVCDBNEPYTARIIPAAGDSSKNITLGEKAKAKNPDG 180
 Db 119 PGSQNSDSTQSVOSTISRFACRIICERNPPTARIYAAGDSSKNITLGEKAKAKMTSDG 178
 QY 181 HMDGLTNGVLVNHPRGFTESQPGVWREISVCGDVYTLRETRSAQORGLVSEETNVL 240
 Db 179 QMDGLTNGVLVNHPRNGFTEDSKPGIMREISVCGVFSLRKTRSAQORGMVEIETNQL 238
 QY 241 QDGLSLDLCGATLLMRTADGLFTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGLSLDLCGATLLMRTAEGLSHTPTVKHLEALROEINARPOCPVGFNTLAFPSMKRKO 298
 QY 301 VVEKOPWATLSCGHVGHYNNWGRSDTEANRECEPCRTVGPYVPLMLGCEAGFYVDAG 360
 Db 299 VVEKOPWVYVNLGCHVGHYNNWGNKBERDCKRECPKRSVGPYVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHFAACPFCAQQLVGEONCIKLIPOGPID 420
 Db 359 PPTHAFSPCGHVSEKKTAWMSQIPLPHGTHFAACPFCAQQLAGEQYIRLIPOGPID 418

RESULT 13

US-10-258-703-4

Sequence 4, Application US/10258703
 Publication No. US20040034199A1
 GENERAL INFORMATION:

APPLICANT: Bird, Timothy A.
 APPLICANT: Coeman, David J.
 TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 FILE REFERENCE: 2990-US
 CURRENT APPLICATION NUMBER: US/10/258,703
 CURRENT FILING DATE: 2002-10-24
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 4
 LENGTH: 418
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-258-703-4

Query Match 83.7%; Score 1917; DB 4; Length 418;
 Best Local Similarity 81.4%; Pred. No. 9,46-180;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGQEHCAPNKEPVYKGLVYLGVNGALPNDGRGRKSRPALYKRPKANGKVPSTVH 60
 Db 1 MFSPDQENH--PSKAPVKYKGLVILVINGSLPNDGRGRKSRPALFKRPKANGKVPSTVH 58
 QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDTDMFQVGRSTESPIDFVYTDI 120
 Db 59 IACTPQAKAISNDQHSISYTLSSRAQTVVVEYTHDSNTDMFQIGRSTESPIDFVYTDV 118
 QY 121 SGSQNTDEAQTOSTISRFACRIVCDBNEPYTARIIPAAGDSSKNITLGEKAKAKNPDG 180
 Db 119 PGSQNSDSTQSVOSTISRFACRIICERNPPTARIYAAGDSSKNITLGEKAKAKMTSDG 178

QY 181 HMDGLTNGVLVNHPRGFTESQPGVWREISVCGDVYTLRETRSAQORGLVSEETNVL 240
 Db 179 QMDGLTNGVLVNHPRNGFTEDSKPGIMREISVCGVFSLRKTRSAQORGMVEIETNQL 238
 QY 241 QDGLSLDLCGATLLMRTADGLFTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGLSLDLCGATLLMRTAEGLSHTPTVKHLEALROEINARPOCPVGFNTLAFPSMKRKO 298
 QY 301 VVEKOPWATLSCGHVGHYNNWGRSDTEANRECEPCRTVGPYVPLMLGCEAGFYVDAG 360
 Db 299 VVEKOPWVYVNLGCHVGHYNNWGNKBERDCKRECPKRSVGPYVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHFAACPFCAQQLVGEONCIKLIPOGPID 420
 Db 359 PPTHAFSPCGHVSEKKTAWMSQIPLPHGTHFAACPFCAQQLAGEQYIRLIPOGPID 418

RESULT 14

US-10-041-030-2

Sequence 2, Application US/10041030
 Publication No. US20020150934A1
 GENERAL INFORMATION:

APPLICANT: Powers, Scott
 APPLICANT: Mu, David
 APPLICANT: Xiang, Phil
 APPLICANT: Peng, Yue
 APPLICANT: Tularik Inc.
 TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
 FILE REFERENCE: 018761-006910US
 CURRENT APPLICATION NUMBER: US/10/041,030
 CURRENT FILING DATE: 2001-12-28
 PRIOR APPLICATION NUMBER: US 60/259,502
 PRIOR FILING DATE: 2001-01-02
 NUMBER OF SEQ ID NOS: 42
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2
 LENGTH: 418
 TYPE: PRT
 ORGANISM: Homo sapiens
 OTHER INFORMATION: human pellino 1
 US-10-041-030-2

Query Match 83.5%; Score 1913; DB 4; Length 418;
 Best Local Similarity 81.4%; Pred. No. 2,36-179;
 Matches 342; Conservative 37; Mismatches 39; Indels 2; Gaps 1;

QY 1 MFSPGQEHCAPNKEPVYKGLVYLGVNGALPNDGRGRKSRPALYKRPKANGKVPSTVH 60
 Db 1 MFSPDQENH--PSKAPVKYKGLVILVINGSLPNDGRGRKSRPALFKRPKANGKVPSTVH 58
 QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDTDMFQVGRSTESPIDFVYTDI 120
 Db 59 IACTPQAKAISNDQHSISYTLSSRAQTVVVEYTHDSNTDMFQIGRSTESPIDFVYTDV 118
 QY 121 SGSQNTDEAQTOSTISRFACRIVCDBNEPYTARIIPAAGDSSKNITLGEKAKAKNPDG 180
 Db 119 PGSQNSDSTQSVOSTISRFACRIICERNPPTARIYAAGDSSKNITLGEKAKAKMTSDG 178
 QY 181 HMDGLTNGVLVNHPRGFTESQPGVWREISVCGDVYTLRETRSAQORGLVSEETNVL 240
 Db 179 QMDGLTNGVLVNHPRNGFTEDSKPGIMREISVCGVFSLRKTRSAQORGMVEIETNQL 238
 QY 241 QDGLSLDLCGATLLMRTADGLFTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGLSLDLCGATLLMRTAEGLSHTPTVKHLEALROEINARPOCPVGFNTLAFPSMKRKO 298
 QY 301 VVEKOPWATLSCGHVGHYNNWGRSDTEANRECEPCRTVGPYVPLMLGCEAGFYVDAG 360
 Db 299 VVEKOPWVYVNLGCHVGHYNNWGNKBERDCKRECPKRSVGPYVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHFAACPFCAQQLVGEONCIKLIPOGPID 420

Db 359 PPTHAFFSCGHVCEKKTATWMSQILPHGHTTFHAACPFCAHQLAGGQYIRLLIFQSPID 418

RESULT 15
US-09-843-905A-2
; Sequence 2, Application US/09843905A
; Patent No. US20020168683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-843-905A-2

Query Match 83.5%; Score 1912; DB 3; Length 418;
Best Local Similarity 81.2%; Pred. No. 2.9e-179;
Matches 341; Conservative 39; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGDEHCAPNKEPVYKGLVLYGNGALPNGDRGRKRSFALYKRPKANGVKPSTVH 60
Db 1 MFSPDQENH--PSKAPVKTGELIVGNGSLPNGDRGRKRSFALYKRPKANGVKPSTVH 58
QY 61 VISTPQASKAISCQGHISITSLRNQTVVVEYTHDKDTMFQVGRSTESPIDFVVTDTI 120
Db 59 IACTPQAKAISNDQHSISYTLSSRAQTVVEYTHDNTDMFQIGRSTESPIDFVVTDTV 118
QY 121 SSGQNTDEAQTQSTIRFACRIYCDNRPYTARIIPAAGDSSKNITFLGEKAAKKNPDG 180
Db 119 PGSQNSDTSVOSTISRFACRIICERSPPTARIYAAGDSSKNITFLGEKAAKMTSDG 178
QY 181 HMDGLTNGVLYMHPRGFTBESQGVWRREISVCGDYVTLRETPSAOQKGLVSEETNVL 240
Db 179 QMDGLTNGVLYMHPRGFTBESQGVWRREISVCGDYVTLRETPSAOQKGLVSEETNVL 238
QY 241 QDGLSLIDLCGATLLMRTADGLFHTPTQKHLEALROEINARPOCPVGIANTLAFPSINRKE 300
Db 239 QDGLSLIDLCGATLLMRTABGLSHPTVYKHLLEALROEINARPOCPVGIANTLAFPSMKRXD 298
QY 301 VVEBKQPMAYLSCGHVGHYNNMGHRSPTENARECEPMCRVTGYPVPLMLGCEAGFYVDAG 360
Db 299 VVEBKQPMVYVINGCHVGHYNNMGHRSPTENARECEPMCRVTGYPVPLMLGCEAGFYVDAG 358
QY 361 PPTHAFFSCGHVCEKKTATWMSQILPHGHTTFHAACPFCAHQLAGGQYIRLLIFQSPID 420
Db 359 PPTHAFFSCGHVCEKKTATWMSQILPHGHTTFHAACPFCAHQLAGGQYIRLLIFQSPID 418

Search completed: March 1, 2006, 21:04:32
Job time: 168 secs

Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 21:00:16 ; Search time 47 Seconds
(without alignments)
738.804 Million cell updates/sec

Title: US-10-041-030-4

Perfect score: 2290

Sequence: 1 MFSPGQEHKAPKKEPVKVG.....ATGVGKQNCIKILPGPID 420

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

- Issued Patents AA:*
- 1: /cgnt2_6/ptodata/1/iaa/5 COMB.pep:*
 - 2: /cgnt2_6/ptodata/1/iaa/6 COMB.pep:*
 - 3: /cgnt2_6/ptodata/1/iaa/7 COMB.pep:*
 - 4: /cgnt2_6/ptodata/1/iaa/8 COMB.pep:*
 - 5: /cgnt2_6/ptodata/1/iaa/9 COMB.pep:*
 - 6: /cgnt2_6/ptodata/1/iaa/10 COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2283	99.7	420	2	US-09-843-905A-8
2	2175.5	95.0	419	2	US-09-843-905A-6
3	1917	83.7	418	2	US-09-843-905A-4
4	1912	83.5	418	2	US-09-843-905A-2
5	1642	71.7	445	2	US-09-843-905A-12
6	1482.5	64.7	406	2	US-10-104-047-3928
7	1302	56.9	424	2	US-09-843-905A-13
8	1235	53.9	455	2	US-09-843-905A-14
9	903	39.4	458	2	US-09-843-905A-15
10	104	4.5	710	2	US-09-252-991A-25700
11	100	4.4	2042	2	US-09-077-098A-6
12	100	4.4	2042	2	US-10-192-584-6
13	96	4.2	717	2	US-09-949-016-6500
14	95.5	4.2	978	2	US-10-198-452A-485
15	95.5	4.2	978	2	US-10-197-220-65
16	95.5	4.2	979	2	US-09-438-185A-456
17	94	4.1	2324	4	PCT-US93-09819-1
18	94	4.1	2324	4	PCT-US93-09819-1
19	93.5	4.1	727	2	US-09-489-039A-14168
20	93.5	4.1	803	2	US-09-543-681A-5399
21	93.5	4.1	1702	2	US-09-854-133-434
22	92	4.0	775	2	US-09-949-016-8030
23	92	4.0	835	2	US-09-949-016-7974
24	90	3.9	2231	1	US-08-153-799-16
25	88.5	3.9	558	1	US-10-037-311B-1
26	88.5	3.9	1104	1	US-08-337-832-5
27	88.5	3.9	1104	1	US-08-828-584-5

28	87.5	3.8	426	2	US-09-252-991A-32215	Sequence 32215, A
29	87	3.8	453	2	US-09-949-016-10252	Sequence 10252, A
30	87	3.8	254	1	US-08-576-626A-32	Sequence 32, Appl
31	86.5	3.8	266	2	US-10-104-047-3891	Sequence 3891, Ap
32	86.5	3.8	498	2	US-09-248-796A-19085	Sequence 19085, A
33	86.5	3.8	1371	2	US-09-902-540-16024	Sequence 16024, A
34	86	3.8	1454	2	US-09-949-016-7599	Sequence 7599, Ap
35	86	3.8	1454	2	US-09-949-016-7600	Sequence 7600, Ap
36	86	3.8	2327	6	5455158-1	Patent No. 5455158
37	86	3.8	2355	2	US-10-360-101-235	Sequence 235, Appl
38	86	3.8	2386	1	US-09-016-366A-12	Sequence 12, Appl
39	86	3.8	2386	2	US-09-961-403-1	Sequence 1, Appl1
40	86	3.8	2446	1	US-08-551-356-2	Sequence 2, Appl1
41	86	3.8	2446	4	PCT-US93-12687-2	Sequence 2, Appl1
42	85.5	3.7	525	1	US-08-160-861-4	Sequence 4, Appl1
43	85.5	3.7	572	1	US-08-160-861-3	Sequence 3, Appl1
44	85.5	3.7	915	2	US-09-817-514A-6	Sequence 6, Appl1
45	85.5	3.7	1073	2	US-09-949-016-9771	Sequence 9771, Ap

ALIGNMENTS

RESULT 1									
US-09-843-905A-8									
Sequence 8, Application US/09843905A									
Patent No. 6703487									
GENERAL INFORMATION:									
APPLICANT: Bird, Timothy A.									
TITLE OR INVENTION: HUMAN PELLINO POLYPEPTIDES									
FILE REFERENCE: 2999A									
CURRENT APPLICATION NUMBER: US/09/843,905A									
CURRENT FILING DATE: 2001-04-27									
PRIOR APPLICATION NUMBER: US 60/200,198									
PRIOR FILING DATE: 2000-04-28									
NUMBER OF SEQ ID NOS: 15									
SOFTWARE: PatentIn version 3.1									
SEQ ID NO 8									
LENGTH: 420									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-09-843-905A-8									
Query Match 99.7%; Score 2283; DB 2; Length 420;									
Best Local Similarity 99.8%; Pred. No. 8.3e-238;									
Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	MFSPGQEHKAPKKEPVKYGELVLYGNGALPAGDRGRKSRPALYKRPKANGKSTVH	60						
DB	1	MFSPGQEHKAPKKEPVKYGELVLYGNGALPAGDRGRKSRPALYKRPKANGKSTVH	60						
QY	61	VISTPQASKAISCKGQHSISYTLRNQTVVETHTKQDMDPFGVSTESPIDFVYTDI	120						
DB	61	VISTPQASKAISCKGQHSISYTLRNQTVVETHTKQDMDPFGVSTESPIDFVYTDI	120						
QY	121	SGSQNTDEAQTOSTISRFACRIVCDRNEPYTARIFAGDSKNIPLFGKAKAKNPDG	180						
DB	121	SGSQNTDEAQTOSTISRFACRIVCDRNEPYTARIFAGDSKNIPLFGKAKAKNPDG	180						
QY	181	HMDGLTTNGVLYVNHPRGFTBESQPGVWREISVCGVYTLRETRSAQQRKLYVESTNV	240						
DB	181	HMDGLTTNGVLYVNHPRGFTBESQPGVWREISVCGVYTLRETRSAQQRKLYVESTNV	240						
QY	241	QDSGLIDLCAGATLMTADGLFTPTQKHIALROEINARPCQGLMTLAPPSINRKE	300						
DB	241	QDSGLIDLCAGATLMTADGLFTPTQKHIALROEINARPCQGLMTLAPPSINRKE	300						
QY	301	VVEKQPMWATLSGCHVGHYNNWGRSDTEANRECEPCMCRTVGPVYVLMGCEAGFYVDAG	360						
DB	301	VVEKQPMWATLSGCHVGHYNNWGRSDTEANRECEPCMCRTVGPVYVLMGCEAGFYVDAG	360						
QY	361	PPHTAFTPCGHVCESEKAKTWSQIPLPHGTHAFAACFPATQVVGKNCIKILPGPID	420						

Db 361 PPTHAFTPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCAATQVGEONCIKILFQGPID 420

RESULT 2

US-09-843-905A-6
Sequence 6, Application US/09843905A

Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 419
TYPE: PRT
ORGANISM: Mus musculus
US-09-843-905A-6

Query Match 95.0%; Score 2175.5; DB 2; Length 419;
Best Local Similarity 95.0%; Pred. No. 3.5e-226;
Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTYH 60
DB 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTYH 60
QY 61 VISTPOAKAISCKGHSISYTLNRQTVVVEYTHKDTPMFGVSTESPIDFVVTDTI 120
DB 61 VISTPOAKAISCKGHSISYTLNRQTVVVEYTHKDTPMFGVSTESPIDFVVTDTI 120
QY 121 SSGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMNPDG 180
DB 121 SSGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMNPDG 180
QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWRISVCGDYTLRETRSAQGRGLVSESTNVL 240
DB 181 HMDGLTTNGVLVNHPRGFTBESOPGVWRISVCGDYTLRETRSAQGRGLVSESTNVL 240
QY 241 QDSGLIDLCGATLLMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAFPSINRKE 300
DB 241 QDSGLIDLCGATLLMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAFPSINRKE 300
QY 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
DB 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
QY 361 PPTHAFTPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCAATQVGEONCIKILFQGPID 420
DB 361 PPTHAFTPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCAATQVGEONCIKILFQGPID 420

RESULT 3

US-09-843-905A-4
Sequence 4, Application US/09843905A

Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1

SEQ ID NO 4
LENGTH: 418
TYPE: PRT
ORGANISM: Homo sapiens
US-09-843-905A-4

Query Match 83.7%; Score 1917; DB 2; Length 418;
Best Local Similarity 81.4%; Pred. No. 3.2e-198;
Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTYH 60
DB 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTYH 60
QY 61 VISTPOAKAISCKGHSISYTLNRQTVVVEYTHKDTPMFGVSTESPIDFVVTDTI 120
DB 61 VISTPOAKAISCKGHSISYTLNRQTVVVEYTHKDTPMFGVSTESPIDFVVTDTI 120
QY 121 SSGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMNPDG 180
DB 121 SSGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMNPDG 180
QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWRISVCGDYTLRETRSAQGRGLVSESTNVL 240
DB 181 HMDGLTTNGVLVNHPRGFTBESOPGVWRISVCGDYTLRETRSAQGRGLVSESTNVL 240
QY 241 QDSGLIDLCGATLLMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAFPSINRKE 300
DB 241 QDSGLIDLCGATLLMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAFPSINRKE 300
QY 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
DB 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
QY 361 PPTHAFTPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCAATQVGEONCIKILFQGPID 420
DB 361 PPTHAFTPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCAATQVGEONCIKILFQGPID 420

RESULT 4

US-09-843-905A-2
Sequence 2, Application US/09843905A

Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 418
TYPE: PRT
ORGANISM: Mus musculus
US-09-843-905A-2

Query Match 83.5%; Score 1912; DB 2; Length 418;
Best Local Similarity 81.2%; Pred. No. 1.1e-197;
Matches 341; Conservative 39; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTYH 60
DB 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTYH 60
QY 61 VISTPOAKAISCKGHSISYTLNRQTVVVEYTHKDTPMFGVSTESPIDFVVTDTI 120
DB 61 VISTPOAKAISCKGHSISYTLNRQTVVVEYTHKDTPMFGVSTESPIDFVVTDTI 120
QY 121 SSGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMNPDG 180
DB 121 SSGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMNPDG 180

```

119 PGSGNSNDTQSVOSTISRFACRIRICERSPPPTAITYAAGPSSKNITLGEBAAMKXSDG 178
181 HMDGLTNGVLVWHPRGFTBESOPGWREISVCGDYVTLRETRSAOQKGLVSESTNVL 240
179 QMGGLTNGVLVWHPRGFTBESOPGWREISVCGDYVTLRETRSAOQKGLVSESTNVL 238
241 QDGLSLIDLCATLMTWRTAGLSTPTVYKHEALROEINAAPOCPVGNLTAPISINRKE 300
239 QDGLSLIDLCATLMTWRTAGLSTPTVYKHEALROEINAAPOCPVGNLTAPISINRKE 298
301 VVEEKOPWATLSCGHVGHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 360
299 VDEKOPWATLSCGHVGHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 358
361 PPTHAFTPCGHVCESEKSAKXWSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 420
359 PPTHAFTPCGHVCESEKSAKXWSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 418

```

```

RESULT 5
US-09-843-905A-12
; Sequence 12, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-843-905A-12

```

```

Query Match 71.7%; Score 1642; DB 2; Length 445;
Best Local Similarity 71.4%; Pred. No. 2e-168;
Matches 299; Conservative 42; Mismatches 74; Indels 4; Gaps 3;

```

```

3 SPOGEHCANKEPVYKGEVLVLYGYNALPNDGRGRKSRPALYKRPKANGVPESTYHVI 62
30 SPG--EDAQGEPIKTKGELIVLYGYNCLASGDKRRSRALSRSHANGVKKPDVWHI 87
63 STPOASFAISCKGQHSISYLSRNQTVVVEYTHDKDTMPQVGRSTESPIDFVVTDTISG 122
88 STPLVSKALSNRQGHSHSYLSRSHSVIYVEYTHDSDTDMFOIGRSTENMIDFVVTDTISG 147
123 SONTDEAQTOSTISRFACRIRICERSPPPTAITYAAGPSSKNITLGEBAAMKXSDG 182
148 G-GAEGSPASQSTISRFACRIRICERSPPPTAITYAAGPSSKNITLGEBAAMKXSDG 206
183 DGLTNGVLVWHPRGFTBESOPGWREISVCGDYVTLRETRSAOQKGLVSESTNVL 242
207 DGLTNGVLVWHPRGFTBESOPGWREISVCGDYVTLRETRSAOQKGLVSESTNVL 266
243 GSLIDLCATLMTWRTAGLSTPTVYKHEALROEINAAPOCPVGNLTAPISINRKE 302
267 GSLIDLCATLMTWRTAGLSTPTVYKHEALROEINAAPOCPVGNLTAPISINRKE 326
303 BEK--QPAVYLSCHVGHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 361
327 PDKOQPMVYVRCGHVGHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 386
362 PTHAFTPCGHVCESEKSAKXWSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 420
387 PSHAFAPCGHVCESEKSAKXWSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 445

```

```

RESULT 6
US-10-104-047-3928
; Sequence 3928, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3928
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3928

```

```

Query Match 64.7%; Score 1482.5; DB 2; Length 406;
Best Local Similarity 65.6%; Pred. No. 3e-151;
Matches 275; Conservative 36; Mismatches 65; Indels 43; Gaps 4;

```

```

3 SPOGEHCANKEPVYKGEVLVLYGYNALPNDGRGRKSRPALYKRPKANGVPESTYHVI 62
30 SPG--EDAQGEPIKTKGELIVLYGYNCLASGDKRRSRALSRSHANGVKKPDVWHI 87
63 STPOASFAISCKGQHSISYLSRNQTVVVEYTHDKDTMPQVGRSTESPIDFVVTDTISG 122
49 STPLVSKALSNRQGHSHSYLSRSHSVIYVEYTHDSDTDMFOIGRSTENMIDFVVTDTISG 108
123 SONTDEAQTOSTISRFACRIRICERSPPPTAITYAAGPSSKNITLGEBAAMKXSDG 182
109 G-GAEGSPASQSTISRFACRIRICERSPPPTAITYAAGPSSKNITLGEBAAMKXSDG 167
183 DGLTNGVLVWHPRGFTBESOPGWREISVCGDYVTLRETRSAOQKGLVSESTNVL 242
168 DGLTNGVLVWHPRGFTBESOPGWREISVCGDYVTLRETRSAOQKGLVSESTNVL 227
243 GSLIDLCATLMTWRTAGLSTPTVYKHEALROEINAAPOCPVGNLTAPISINRKE 302
228 GSLIDLCATLMTWRTAGLSTPTVYKHEALROEINAAPOCPVGNLTAPISINRKE 287
303 BEK--QPAVYLSCHVGHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 361
288 PDKOQPMVYVRCGHVGHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 347
362 PTHAFTPCGHVCESEKSAKXWSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 420
348 PSHAFAPCGHVCESEKSAKXWSQIPLPHGTHAFHAACPCATOLVGEONCIKLIPOGPID 406

```

```

RESULT 7
US-09-843-905A-13
; Sequence 13, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Drosophila melanogaster

```


RESULT 10
US-09-252-991A-25700
Sequence 25700, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 31142
SEQ ID NO 25700
LENGTH: 710
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25700

Query Match 4.5%; Score 104; DB 2; Length 710;
Best Local Similarity 22.3%; Pred. No. 0.1;
Matches 90; Conservative 44; Mismatches 162; Indels 108; Gaps 21;

QY 29 GALEPNDGRGRKSRFALYKRPKANGKSTVTVISTPOAKAIS-----CKGQ 76
DB 264 GLQPDARRRRAAVIGF-RARRGAPPRIVIGTMOGAMSELITDIAVIGGAAGI 322
QY 77 HSIISYLSRQTVVVEYTHDKDTPMFOYGSSTESPIDPVYDTITSSQNTDEAOITSTI 136
DB 323 GVIALSLKRSPGLISLIEPADTHYQPGWTLVGGAQAQDITD-----RPMGLVPPGV 377
QY 137 SRPACRIVCRNREYTAIRAPAGFDS--SKNIPL--GEKAKMKNPQMDGLTNTGV-- 190
DB 378 EMATRY--RYVDPEARLLLEGDSLEYRNLIVCPGLRLA-WRISGLEETLRNAYTS 434
QY 191 -----LVMPHGG--FTESQPGWREISVCGDVYTLRETRSAOQKGLVES 235
DB 435 NYRYDLAPYTWELVRGKGLAFTGPAMP-----ICAG-----APQKAWYLSLC 479
QY 236 E---TNVLDDGSL-IDLCGATILMRADGIFHTPTQKHIEALROEINARPOCPVELANT 290
DB 480 DHMLRECVLDDIEVEFDLAGAA-LFGVAD--FVPLMEYVRYKYSABE-----AFNSNL 529
QY 291 LAFPSIRKEVEKQWAVYLSGHHVGHNMGRSDTEAN---ERECPMCRVTGPPYVP 346
DB 530 VKYDGAARKA-----WFEVKADGNTNLAEKDFDLHVVPPQLP 568
QY 347 -----LMLGCEAGF-YVDAGPPTHA---FTPCGHVCESEKSAK 379
DB 569 PTFFAASGLDAGMCEVDPATLQHVHGEIFALGDVCGTANAK 612

RESULT 11
US-09-077-098A-6
Sequence 6, Application US/09077098A
Patent No. 6544519
GENERAL INFORMATION:
APPLICANT: TOKUNAGA, Ei-ji
SAKAGUCHI, Masaashi
MATSUO, Kazuo
HAMADA, Fukuaburo
TOKIYOSHI, Sachio
TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 624 Ninth Street, N.W., Suite 300
CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077.098A
FILING DATE: 19-May-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP97/03222
FILING DATE: 12-SEP-1997
APPLICATION NUMBER: JP 27,148/1996
FILING DATE: 19-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: KORNBAU, Anne M.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TOKUNAGA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2042 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-077-098A-6

Query Match 4.4%; Score 100; DB 2; Length 2042;
Best Local Similarity 22.7%; Pred. No. 1.6;
Matches 73; Conservative 45; Mismatches 124; Indels 80; Gaps 18;

QY 53 GKVPSTVHYVISTP--QASRAISCKGQHSIS---YTLNRQTVVVEYTHDKD-TDMFOVG 105
DB 1346 GLDDTYTLINKNNPADDLNLSLSESGKNAITGLVDVVKTKNSPTVPSPTDSNKKKFTVG 1405
QY 106 RSTESPDPFVPTPTTISGQNTDEAOITQS-----TISRACRIVCD-----R 147
DB 1406 -----VDF--DTITTEGATDNDKLTITSKVSYSYTNKLANFTDILSDGSGNAATTA 1457
QY 148 NEPYTAIRIFAPGD--SSKNIFLGEK-----AAKKNPDG-----HMDGLTIN--GV 190
DB 1458 NDGVGRRLSDGFTISENPTLGSKQYNGSDSLGVWYDQNGVFKLSINMTALTLSILANT 1517
QY 191 LVNHPRGFTESQPGWREISVCGDVYTLRETRSAOQKGLVESFTNVL-----QDGS- 244
DB 1518 FAKLDASNNLTDSDNKKWR---TALNVYSKTEVDAREIKSKVTLTTPDSGLIFATKQAGSG 1574
QY 245 ---LIDLCGATILMRADGIFHTPT-----OKHIEALROEINARPOCPVG 287
DB 1575 NNAGID-AGKKKISNVADGDI-SPTSGDVYVTRGQLYALMOKGIRVYGDEVSPKTYTAP 1632
QY 288 LNTLAFPSIRKEVEKQWPA 309
DB 1633 TN--ANPTATAPTASTQGMA 1652

RESULT 12
US-10-192-584-6
Sequence 6, Application US/10192584
Patent No. 6918080
GENERAL INFORMATION:
APPLICANT: TOKUNAGA, Ei-ji
SAKAGUCHI, Masaashi
MATSUO, Kazuo
HAMADA, Fukuaburo
TOKIYOSHI, Sachio
TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS

PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME

NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 624 Ninth Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/192,584
FILING DATE: 11-Jul-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/077,098
FILING DATE: 19-May-1998
APPLICATION NUMBER: PCT/JP97/03222
FILING DATE: 12-SEP-1997
APPLICATION NUMBER: JP 27,148/1996
FILING DATE: 19-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: KORNBAU, Anne M.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TOKUNAGA-1
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2042 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-192-584-6

Query Match 4.4%; Score 100; DB 2; Length 2042;
Best Local Similarity 22.7%; Pred. No. 1.6;
Matches 73; Conservative 45; Mismatches 124; Indels 80; Gaps 18;

53 GVKPSTVHVISTP-QASKAISCGGHSIS---YTSRQGVVVEYTHDKD-TDMQVG 105
1346 GLDDTTLINKINPADQLSNLSESGKNAITGLVDVVKTKNSPIVPESTDSNKKQFTVG 1405
106 RSTESPIDFVVTITSGSONTDEAQTOS-----TISRFACTIVCD-----R 147
1406 -----VDF--TDITIGDATDDKGLTTTSKSVESYTNKLANPSTDLILDGSGNATTA 1457
148 NEPYTARIFAAGPD-SKKNIFLGEK-----AAKKNPDG-----HMDGLTTN--GV 190
1458 NDGVGKRRLLSDGFIKSENFILSGKQNGSDSLGVMDDQNGVFKLSLMTALITTSANT 1517
191 LVNHPRGFTBESGPGVWREISVCGDYVTLRETRSAOQKGLVSEINVL-----QDGS- 244
1518 PAKLDASNLTDSDSKKEMR--TALNYSKTEVDAAEQKSKVTLTPPSGLIFATKQAGSG 1574
245 ---LIDCGATLWRTDGLFHTPT-----QKHEALRQEIINARPOCPVG 287
1575 NMAGID-AGNKKISNVADGDI-SPTSGDVVTGRQLYALMOKGIRVYGDVSPYTKQTTPAR 1632
288 LNTLAPPSINRKEVEEKOPWA 309
1633 TN--ANPTATTATPASTQGWA 1652

RESULT 13
US-09-949-016-6500
Sequence 6500, Application US/09949016

Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: PasteSeq for Windows Version 4.0
SEQ ID NO 6500
LENGTH: 717
TYPE: PRT
ORGANISM: Human
US-09-949-016-6500

Query Match 4.2%; Score 96; DB 2; Length 717;
Best Local Similarity 22.4%; Pred. No. 0.76;
Matches 71; Conservative 44; Mismatches 110; Indels 92; Gaps 16;

52 NGVKPSTVHVISTPQASKAISCG-----KGHSISYTSRQGVVVEYTHDKDQDMFOYG 105
211 DSVTLNTHVIGGFDRAVTCIAPSKNSGNGNLCAVDSDNDHVLVWMOKEBKADYK 270
106 RSTES--PIDFVVTD---ISGSONTDEAQTOSTI-----SRFACRIVC-----D 146
271 CSNEAVFADPHRTDNIIVTQGEITSLDPRRKPLIRSKDYSTRAKAVLCTVPE 330
147 RNEPYTARIFAAGPDSSKNIFL---GEKAKKNPDGMDGLT-----TNGVLVNHPRG 198
331 NGDTITG-----DSSGNILWKGKGTNRISYAVQGAHGGISPLCLMDGLTV---SGG 380
199 FTBESGPGVWREISVCGDYVTLRETRSAOQKGL---VSEINVLDDGSLDLCGAT--- 252
381 GKOR-----KLISWGNQKLNKTEIPPOGPIRTVAGKGDV1-----LIGTRNF 427
253 LWRITDGLFHTPTQKIBAL-ROEINARPOCPVGLNTLAPPSINRKEVEEKOPWAYL 311
428 VLQGTISGDPFTPTTQGTDELMGLAIHASKPQ-----PL 461
312 SCGHVGHYHNW---GHR 325
462 TCGHDKHATLMDAVGHR 478

RESULT 14
US-09-198-452A-485
Sequence 485, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Grifols, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 485
LENGTH: 492
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-485

Query Match 4.2%; Score 95.5; DB 2; Length 492;
Best Local Similarity 22.1%; Pred. No. 0.45;
Matches 47; Conservative 25; Mismatches 74; Indels 67; Gaps 9;

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 20:53:05 ; Search time 232 Seconds
(without alignments)
1277.250 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MFSPGQEHKAPKPKYKVG.....ATQVGEQNCIKLIFGCPID 420

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2290	100.0	420	1	PEL12_HUMAN
2	2172.5	94.9	419	1	PEL12_MOUSE
3	2138.5	93.4	419	2	06G057_XENLA
4	2129.5	93.0	419	2	06P4Y5_XENTR
5	2006	87.6	428	2	06DHJ8_BRARE
6	1917	83.7	418	1	PEL11_HUMAN
7	1917	83.7	418	2	053176_HUMAN
8	1912	83.5	418	1	PEL11_MOUSE
9	1912	83.5	418	2	05SRW7_MOUSE
10	1884	82.3	418	2	07ZXU3_XENLA
11	1864	81.4	405	2	05ZKT7_CHICK
12	1788	78.1	428	2	04SGC9_TETNG
13	1647	71.9	440	2	04SEB1_TETNG
14	1632	71.3	445	1	PEL13_MOUSE
15	1620	70.7	469	1	PEL13_HUMAN
16	1493	65.2	420	1	04SGC7_TETNG
17	1302	56.9	424	1	PEL11_MOUSE
18	1279.5	55.9	389	2	07POM0_ANOGA
19	1235	53.9	455	2	09NDP9_CIOIN
20	1100	48.0	441	2	04RLZ6_TETNG
21	910	39.7	458	2	022967_CABEL
22	903	39.4	458	2	09NAP3_CABEL
23	883.5	38.6	450	2	061KK0_CABER
24	602	26.3	121	2	0562B8_RAT
25	446.5	19.5	146	2	06Q430_DROBU
26	277	12.1	49	2	0659D8_HUMAN
27	264	11.5	75	2	05ZM03_CHICK
28	212	9.3	57	2	04RAV6_TETNG
29	122.5	5.3	215	2	09YVJ8_MSEPV
30	119	5.2	340	2	08SVX5_ENCCU
31	118	5.2	1358	2	07QCO6_ANOGA

32	116.5	5.1	1084	2	06NK13_CORDI
33	113	4.9	804	2	07TPT4_MOUSE
34	113	4.9	818	2	08CC59_MOUSE
35	113	4.9	818	2	09DBC8_MOUSE
36	105.5	4.6	1060	2	059GK8_HUMAN
37	103.5	4.5	1591	1	TIAM1_MOUSE
38	103	4.5	675	2	07QXL4_GIALA
39	102.5	4.5	1857	2	08TW55_METAC
40	102	4.5	620	2	096DH7_HUMAN
41	102	4.5	689	2	09URU4_MOUSE
42	102	4.5	710	2	06PB38_MOUSE
43	102	4.5	711	2	08BQC8_MOUSE
44	102	4.5	815	2	08WVL6_HUMAN
45	101.5	4.4	900	2	06PID6_MOUSE

ALIGNMENTS

RESULT 1	PEL12_HUMAN	STANDARD;	PRT;	420 AA.
AC	Q9HAT8;			
AD	29-MAR-2004 (Rel. 43, Created)			
DT	29-MAR-2004 (Rel. 43, Last sequence update)			
DT	10-MAY-2005 (Rel. 47, Last annotation update)			
DE	Pellino protein homolog 2 (Pellino 2).			
GN	Name=PEL12;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE=21203570; PubMed=11306823;			
RA	Resch K., Jockusch H., Schmitt-John T.;			
RT	"Assignment of homologous genes, Pel11/Pellino and Pel12/Pellino, for the			
RT	Pelle adaptor protein Pel11 to mouse chromosomes 11 and 14 and human			
RT	chromosomes 2p13.3 and 14q21, respectively, by physical and radiation			
RT	hybrid mapping."			
RL	Cytogenet. Cell Genet. 92:172-174(2001).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	TISSUE=Lung;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.;			
RA	Klausner R.D., Collins P.S., Wagner L., Shennan C.M., Schuler G.D.;			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.;			
RA	Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Heien F.;			
RA	Diatchenko L., Markovina K., Farmer A.A., Rubin G.M., Hong L.;			
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.;			
RA	Brownstein M.J., Urdin T.B., Tohiyuki S., Carninci P., Prange C.;			
RA	Raba S.S., Longellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.;			
RA	Beak S.A., McMan P.J., McKernan K.J., Malek J.A., Gamarane P.H.;			
RA	Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.;			
RA	Villalón D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.;			
RA	Fahy J., Hailton E., Kettman M., Madan A., Rodriguez S., Sanchez A.;			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.;			
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.;			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.;			
RA	Butterfield Y.S.N., Krzywnicki M.T., Skalska U., Smalins D.E.;			
RA	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[3]			
RP	FUNCTION AND INTERACTIONS WITH TRAP6 AND MAP3K7.			
RX	MEDLINE=22689054; PubMed=12804775; DOI=10.1016/S0014-5793(03)00533-7;			
RA	Jensen L.E., Whitehead A.S.;			
RT	"Pellino2 activates the mitogen activated protein kinase pathway."			
RL	FEBS Lett. 545:199-202(2003).			
RN	[4]			

RP FUNCTION, PHOSPHORYLATION, AND INTERACTIONS WITH IRAK1 AND IRAK4.
 RX MEDLINE=22744764; PubMed=12860405; DOI=10.1016/S0014-5793(03)00697-5;
 RA Strelow A., Kollwe C., Wesche H.;
 RT "Characterization of Pellino2, a substrate of IRAK1 and IRAK4.";
 RL FEBS Lett. 547:157-161(2003).
 CC - FUNCTION: Scaffold protein which probably links Toll-like
 CC receptors (TLRs) to basic cellular processes via its interaction
 CC with the complex containing IRAK kinases and TRAF6. Can activate
 CC the MAP (mitogen activated protein) kinase pathway leading to
 CC activation of ERK1. Not required for NF-kappa-B activation.
 CC - SUBUNIT: Interacts with TRAF6, IRAK1, IRAK4 and MAP3K7.
 CC - INTERACTION:
 CC P51617:IRAK1; NbExp=1; Intact=EBI-448407, EBI-358664;
 CC Q9NMW3:IRAK4; NbExp=1; Intact=EBI-448407, EBI-448378;
 CC - PTM: Phosphorylated by IRAK1 and IRAK4.
 CC - SIMILARITY: Belongs to the pellino family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AF302502; AAG15390.1; -; mRNA.
 CC EMBL; BC009476; AAH09476.1; -; mRNA.
 CC Intact; Q9HAT8; -
 CC DR Ensembl; ENSG00000139946; Homo sapiens.
 CC DR HGNC; HGNC:8828; PELI2.
 CC DR H-InvDB; HIK0011690; -
 CC DR InterPro; IPR006800; Pellino.
 CC DR PANTHER; PTHR12098; Pellino; 1.
 CC DR Pfam; PF04710; Pellino; 1.
 CC KM Phosphorylation.
 SO SEQUENCE 420 AA; 46435 MW; 2FC5E661C13BC11A CRC64;

Query Match 100.0%; Score 2290; DB 1; Length 420;
 Best Local Similarity 100.0%; Pred. No. 1.5e-182;
 Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESPGEHCAKPEPKYGEIVLVLYGNGALPNDRRRRRSPRLYKPKANGKYSTYH 60
 DB 1 MESPGEHCAKPEPKYGEIVLVLYGNGALPNDRRRRRSPRLYKPKANGKYSTYH 60
 QY 61 VISTPQASKAISKQGHISYTLSSRNQTVVEYHDTQTMFOYGRSTESPDPVVDTI 120
 DB 61 VISTPQASKAISKQGHISYTLSSRNQTVVEYHDTQTMFOYGRSTESPDPVVDTI 120
 QY 121 SSGQNTDEAQITOSTISRFACRIYCDNRPYTAIFAAGFDSKNIFLGEKAAKXNDG 180
 DB 121 SSGQNTDEAQITOSTISRFACRIYCDNRPYTAIFAAGFDSKNIFLGEKAAKXNDG 180
 QY 181 HMGGLTNGVLNHPRGFTBESQGVWRISVCGDYTLRFRSAOQRKLVSESNVL 240
 DB 181 HMGGLTNGVLNHPRGFTBESQGVWRISVCGDYTLRFRSAOQRKLVSESNVL 240
 QY 241 QDSGLDLCGATLMRTADGLFHTPTQKHIEALROEINAAPOCPVGLNTLAPSTIRKE 300
 DB 241 QDSGLDLCGATLMRTADGLFHTPTQKHIEALROEINAAPOCPVGLNTLAPSTIRKE 300
 QY 301 VVEEKQFWAYLSCGHVGYNHWGRSDTEANERECPCMCRTVGPVPLMLCEAGFYVDAG 360
 DB 301 VVEEKQFWAYLSCGHVGYNHWGRSDTEANERECPCMCRTVGPVPLMLCEAGFYVDAG 360
 QY 361 PPTHAFPCGHVSEKSAKXWSQPLPHGTHAFPAACPFCAQTQVGEONCIKLIPOCPID 420
 DB 361 PPTHAFPCGHVSEKSAKXWSQPLPHGTHAFPAACPFCAQTQVGEONCIKLIPOCPID 420

DT 29-MAR-2004 (Rel. 43, last sequence update)
 DT 10-MAY-2005 (Rel. 47, last annotation update)
 DE Pellino protein homolog 2 (Pellino 2).
 GN Name=Peli2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RC STRAIN=C57BL/6J;
 RX MEDLINE=21203570; PubMed=11306823;
 RA Resch K., Jockusch H., Schmitt-John T.;
 RT "Assignment of homologous genes, Peli1/Pell1 and Peli2/Pell2, for the
 RT Peli adaptor protein Pellino to mouse chromosomes 11 and 14 and human
 RT chromosomes 2p13.3 and 14q21, respectively, by physical and radiation
 RT hybrid mapping.";
 RL Cytogenet. Cell Genet. 92:172-174(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA) (ISOFORMS 1 AND 3).
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum, Epididymis, and Pituitary;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nakai I., Osato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schirml L.M., Kampin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Brad D., Brusic V., Chothia C., Corbett L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser J.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson J.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglocz D.R., Maltais L., Marchionni L., McKenna B., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perle G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi A., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Varadachari R., Wagner L., Whitehead C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boriss A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kikukawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shidaka K., Shingawa A.,
 RA Yasunishi A., Yoshino Y., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA) (ISOFORM 2).
 RC TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshuler S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Kana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
 RA Villilion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodriguez Y., Bouffard G.G.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences."
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[4]
FN	FUNCTION, TISSUE SPECIFICITY, AND INTERACTION WITH IRAK1.
RX	MEDLINE=22257640; PubMed=12370331;
RA	Yu K.-Y., Kwon H.-J., Norman D.A.M., Vig E., Goebel M.G.,
RA	Harrington M.A.;
RT	"Mouse pellino-2 modulates IL-1 and lipopolysaccharide signaling.";
RL	J. Immunol. 169:4075-4078(2002).
CC	-1- FUNCTION: SCF fold protein which probably links Toll-like
CC	receptors (TLRs) to basic cellular processes via its interaction
CC	with the complex containing IRAK kinases and TRAF6. Can activate
CC	the MAP (mitogen activated protein) kinase pathway leading to
CC	activation of Elk1. Not required for NF-kappa-B activation.
CC	-1- SUBUNIT: Interacts with TRAF6, IRAK4 and NABP37 (By similarity).
CC	Interacts with IRAK1.
CC	-1- INTERACTION:
CC	OS2406:Irk1; NBExp=1; IntAct=EBI-448554, EBI-448533;
CC	-1- ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=3;
CC	Name=1;
CC	IsoId=Q8BST6-1; Sequence=Displayed;
CC	Name=2;
CC	IsoId=Q8BST6-2; Sequence=VSP_008636;
CC	Note=No experimental confirmation available;
CC	Name=3;
CC	IsoId=Q8BST6-3; Sequence=VSP_008637, VSP_008638;
CC	Note=No experimental confirmation available;
CC	-1- TISSUE SPECIFICITY: Widely expressed both in embryos and adult.
CC	Weakly or not expressed in spleen and thymus.
CC	-1- PTM: Phosphorylated by IRAK1 and IRAK4 (by similarity).
CC	-1- SIMILARITY: Belongs to the pellino family.
CC	-----
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use as long as its content is in no way modified and this statement is not
CC	removed.
CC	-----
DR	EMBL; AF302504; AAG15392.1; -; mRNA.
DR	EMBL; AK030564; BAC27024.1; -; mRNA.
DR	EMBL; AK033815; BAC28485.1; -; mRNA.
DR	EMBL; AK082342; BAC38472.1; -; mRNA.
DR	EMBL; BC027062; AAH27062.1; -; mRNA.
DR	IntAcC; Q8BST6; -
DR	Ensembl; ENSMUSG0000021846; Mus musculus.
DR	MGI; MGI:1891445; Pel12.
DR	GO; GO:000515; F:protein binding; IPI.
DR	InterPro; IPRO06800; Pellino.
DR	PANTHER; PTHR12098; Pellino.1.
DR	Pfam; PF04710; Pellino; 1.
KW	Alternative splicing; Phosphorylation.
FT	VASPLIC
FT	1
FT	100
FT	Missing (in isoform 2).
FT	/ftid=vsp_008636.
FT	AISNGHSISTYLSRSQTVVVEYTHDKDTMPGVGRSTES
FT	PIDFVLTVSGGQNEADQIOSTISRACIVCRNPYT
FT	ARIF -> LIPAKHYNEADSESLATLKRDFLTGCSSO
FT	RRRYDPAFRGGSGSNQVAQAALICPLSYTIKKQDIR
FT	CLKIKTFSCMS (in isoform 3).
FT	/ftid=vsp_008637.
FT	Missing (in isoform 3).
FT	/ftid=vsp_008638.
FT	E -> K (in Ref. 2; BAC27024).
FT	A -> T (in Ref. 2; BAC27024).
FT	R -> G (in Ref. 2; BAC38472).
FT	RR -> KK (in Ref. 1).
FT	A -> T (in Ref. 1).
FT	S -> H (in Ref. 1).
FT	T -> A (in Ref. 3).
FT	R -> M (in Ref. 1).
FT	A -> V (in Ref. 1).
FT	SEQUENCE
FT	419 AA; 46272 MW; 786C92C286C3B0DCB CRC64;
VS	VARSPLIC
VS	156
VS	419
VS	CONFICT
VS	7
VS	7
VS	11
VS	20
VS	20
VS	38
VS	39
VS	44
VS	320
VS	327
VS	327
VS	338
VS	338
VS	364
VS	364
VS	419 AA; 46272 MW; 786C92C286C3B0DCB CRC64;

Query Match	94.9%	Score 2172.5	DB 1	Length 419
Best Local Similarity	95.0%	Pred. 0.172-173		
Matches 399	Conservative 10	Mismatches 10	Indels 1	Gaps 1
QY	1	MFSPQEEHCAPNTEPVKYGELVVLGYNGALPNDRGRKSRFPALYKRPKANGVKESFTVH	60	
Db	1	MFSPQEEHSPSPNPKPVYRELVLGYNGALPNDRGRKSRFPALYKRTTVASGVSPSTIH	60	
QY	61	VISTPQASKAISCKGHSISTVLSRNOVTVVEYTHDQTMDFQVGRSTESPIDFVVTPTI	120	
Db	61	WVTPQASKAISCKGHSISTVLSRNOVTVVEYTHDQTMDFQVGRSTESPIDFVVTPTI	120	
QY	121	SGSONTDEAQTOSTIRSFACRIYCDNENPEYTAIFIAGPSSKNIFLGEFAAKKPNBDG	180	
Db	121	SGGQNEB-AQTOSTIRSFACRIYCDNENPEYTAIFIAGPSSKNIFLGEFAAKKPNBDG	179	
QY	181	HMDGLTTNGVILVMPHRCGFTEESOPGWREISVCGDYVTLTETRSAAQRGULVSEETVL	240	
Db	180	HMDGLTTNGVILVMPHRCGFTEESOPGWREISVCGDYVTLTETRSAAQRGULVSEETVL	239	
QY	241	ODGSLIDLCGATLLMRPADGLFHTPTQKTHIALOEINAAAPQCPVGLNTLAPSIINKE	300	
Db	240	QDGSLLIDLCGATLLMRPADGLFHPAPQKTHIALOEINAAAPQCPVGLNTLAPSIINKE	299	
QY	301	VVEKQKPWAVYSCGHVGYHNHGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG	360	
Db	300	VVEKQKPWAVYSCGHVGYHNHGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG	359	
Db	360	PPTHAFTPCGHVCESEKSAKTSQIPLPHGTHAFPAACFCATQVGBONCIKTLFQGEVD	419	
RESULT 3				
Q6G057_XENTIA		PRT;	419 AA.	
ID	Q6G057_XENTIA	PRELIMINARY;		
AC	Q6G057;			
DT	05-JUL-2004 (TREMBlrel. 27, Created)			
DT	05-JUL-2004 (TREMBlrel. 27, Last sequence update)			
DT	05-JUL-2004 (TREMBlrel. 27, Last annotation update)			
DE	MG80529 ProteIn.			
GN	Name=MG80329;			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;			
OC	Xenopodinae; Xenopus; Xenopus.			
OX	NCBI_TaxID=8335;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Ovary;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shemem C.M., Schuler G.D.,			
RA	Aleschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Wathey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Butterfield Y.S.N., Kiryakin M.I., Skalski V., Smalins D.E.,			
RA	Scherer A., Schein J.E., Jones S.J.M., Marra M.A.,			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RL	and mouse cDNA sequences".			
RN	Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903 (2002).			
	[2]			

RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Ovary;
 RX MEDLINE=2234113; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strusberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT Initiative";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Ovary;
 RA Klein S., Gerhard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC072891; AA472891.1; -; mRNA.
 DR Interpro: IPR006800; Pellino.
 DR Pfam: PF04710; Pellino; 1.
 DR SEQUENCE 419 AA; 46372 MW; 55AC2FE7953D6145 CRC64;
 SO
 Query Match 93.4%; Score 2138.5; DB 2; Length 419;
 Best Local Similarity 91.2%; Pred. No. 6.6e-170;
 Matches 383; Conservative 24; Mismatches 12; Indels 1; Gaps 1;
 QY 1 MFSPGQEHKCAPKPEPVYKGLVVLGYNGALPNDGRGRKSRFPALYKRPANGVKSTVH 60
 DB 1 MFSPGQEHKCAPKPEPVYKGLVVLGYNGCLPNDGRGRKSRFPALYKRPANGVKSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLRNQTVVVEYTHDKDTMFQVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLRNQTVVVEYTHDKDTMFQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRACRIVCDRNPYTRARIFAAGPSSKNIFLGEKAKMKNPDG 180
 DB 121 SGNQN-DETQITQSTISRACRIVCDRNPYTRARIFAAGPSSKNIFLGEKAKMKNPDG 179
 QY 181 HMDGLTNGVLVNHPRGFTBESOPGWREISVCGDVYTLRETRSAQGRGLVESETNVL 240
 DB 180 HMDGLTNGVLVNHPRGFTBESOPGWREISVCGDVYTLRETRSAQGRGLVESETNVL 239
 QY 241 QDGSLLDLCGATLMTADGLFTPTQKHIALROEINARPCQVGLNTLAFPSINRKE 300
 DB 240 QDGSLLDLCGATLMTADGLFTPTQKHIALROEINARPCQVGLNTLAFPSINRKE 299
 QY 301 VVEKQPMAYLSCGHVGHYHNGHRSDEANERECPCRTGVPYVPLMLGCEAGFYVDAG 360
 DB 300 VVEKQPMAYLSCGHVGHYHNGHRSDEANERECPCRTGVPYVPLMLGCEAGFYVDAG 359
 QY 361 PPTHAFPCGHVSEKSAKYSQIPLPHGTHAFHACPCATQVSEONCICKIFQGPID 420
 DB 360 PPTHAFPCGHVSEKSAKYSQIPLPHGTHAFHACPCATQVSEONCICKIFQGPID 419
 RESULT 4
 Q6P4Y5_XENTR PRELIMINARY; PRT; 419 AA.
 ID Q6P4Y5_XENTR PRELIMINARY; PRT; 419 AA.
 AC Q6P4Y5;
 DT 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
 DE 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 GN Hypothetical protein MGC75650.
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus; Silurana.
 NC NCB1_faxid=8364;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strusberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
 RA Krausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Butow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
 RA Brownstein M.J., Ustin T.B., Tothiyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mollath S.J.,
 RA Bobak S.A., McSwan P.J., McKernan K.J., Males J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerich A., Schein J.E., Jones S.J.M., Maria M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC063200; AA463200.1; -; mRNA.
 DR Interpro: IPR006800; Pellino.
 DR Pfam: PF04710; Pellino; 1.
 DR Hypothetical protein.
 DR SEQUENCE 419 AA; 46328 MW; 042A50786CE69D75 CRC64;
 SO
 Query Match 93.0%; Score 2129.5; DB 2; Length 419;
 Best Local Similarity 91.4%; Pred. No. 3.7e-169;
 Matches 384; Conservative 22; Mismatches 13; Indels 1; Gaps 1;
 QY 1 MFSPGQEHKCAPKPEPVYKGLVVLGYNGALPNDGRGRKSRFPALYKRPANGVKSTVH 60
 DB 1 MFSPGQEHKCAPKPEPVYKGLVVLGYNGCLPNDGRGRKSRFPALYKRPANGVKSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLRNQTVVVEYTHDKDTMFQVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLRNQTVVVEYTHDKDTMFQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRACRIVCDRNPYTRARIFAAGPSSKNIFLGEKAKMKNPDG 180
 DB 121 SGNQN-DETQITQSTISRACRIVCDRNPYTRARIFAAGPSSKNIFLGEKAKMKNPDG 179
 QY 181 HMDGLTNGVLVNHPRGFTBESOPGWREISVCGDVYTLRETRSAQGRGLVESETNVL 240
 DB 180 HMDGLTNGVLVNHPRGFTBESOPGWREISVCGDVYTLRETRSAQGRGLVESETNVL 239
 QY 241 QDGSLLDLCGATLMTADGLFTPTQKHIALROEINARPCQVGLNTLAFPSINRKE 300
 DB 240 QDGSLLDLCGATLMTADGLFTPTQKHIALROEINARPCQVGLNTLAFPSINRKE 299
 QY 301 VVEKQPMAYLSCGHVGHYHNGHRSDEANERECPCRTGVPYVPLMLGCEAGFYVDAG 360
 DB 300 VVEKQPMAYLSCGHVGHYHNGHRSDEANERECPCRTGVPYVPLMLGCEAGFYVDAG 359
 QY 361 PPTHAFPCGHVSEKSAKYSQIPLPHGTHAFHACPCATQVSEONCICKIFQGPID 420
 DB 360 PPTHAFPCGHVSEKSAKYSQIPLPHGTHAFHACPCATQVSEONCICKIFQGPID 419
 RESULT 5
 Q6DHU8_BRARE PRELIMINARY; PRT; 428 AA.
 ID Q6DHU8_BRARE PRELIMINARY; PRT; 428 AA.
 AC Q6DHU8;
 DT 25-OCT-2004 (Tremblrel. 28, Created)
 DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
 DE 25-OCT-2004 (Tremblrel. 28, Last annotation update)
 GN Name=pe12;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.

NCBI_TaxID=7955;
[1]
NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schmech A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RG NIH MGC Project;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC075973; AAH75973.1; -; mRNA
DR ZFIN: ZDB-GENE-040718-360; zgc:92268.
DR InterPro: IPR006800; Pellino.
DR Pfam: PF04710; Pellino; 1.
SQ SEQUENCE 428 AA; 47287 MW; B33890B49F4366A CRC64;

Query Match 87.6%; Score 2006; DB 2; Length 428;
Best Local Similarity 84.3%; Pred. No. 8.1e-159;
Matches 361; Conservative 28; Mismatches 31; Indels 8; Gaps 2;

1 MFSPGQGEHCAPNPKYKGLVVLGYNGALPNDGRKRKSPALYKRPANGVSPSTVH 60
1 MFSPGQGEHCAPNPKYKGLVVLGYNGALPNDGRKRKSPALYKRPANGVSPSTVH 60
61 VISTPQASKAISCKQHSISYTSRNQTVVVEYTHDKOTDMPQVGRSTESPIDFVDTI 120
61 VISTPQASKAISCKQHSISYTSRNQTVVVEYTHDKOTDMPQVGRSTESPIDFVDTI 120
121 SSGSNDDEAQTOSTIRPACRIYCDNREPTTAIPAAGPDSKNIFLGEKAAKKNPDG 180
121 SSGSNDDEAQTOSTIRPACRIYCDNREPTTAIPAAGPDSKNIFLGEKAAKKNPDG 180
121 SSGSESEETITOSTIRPACRIYCDNREPTTAIPAAGPDSKNIFLGEKAAKKNPDG 180
121 SSGSESEETITOSTIRPACRIYCDNREPTTAIPAAGPDSKNIFLGEKAAKKNPDG 180
181 HMDGLTNGVLMVMPRGSGFTSEESQGVWRREISVCGDVYTLRETRSAORGKLVSESTNVL 240
181 HMDGLTNGVLMVMPRGSGFTSEESQGVWRREISVCGDVYTLRETRSAORGKLVSESTNVL 240
181 HMDGLTNGVLMVMPRGSGFTSEESQGVWRREISVCGDVYTLRETRSAORGKLVSESTNVL 240
241 QDSGLIDLCAGATILMRTADGLFHTPTQKHEALROEINARPOCPVGLANTLAPPSINR-- 298
241 QDSGLIDLCAGATILMRTADGLFHTPTQKHEALROEINARPOCPVGLANTLAPPSINR-- 298
241 QDSGLIDLCAGATILMRTADGLFHTPTQKHEALROEINARPOCPVGLANTLAPPSINR-- 298
241 QDSGLIDLCAGATILMRTADGLFHTPTQKHEALROEINARPOCPVGLANTLAPPSINR-- 298
299 -----KEVSEKQPMATLSCGHVGYNNNGHRSSTEAN-ERECMGTCTGPPVPLMTGCE 352
299 -----KEVSEKQPMATLSCGHVGYNNNGHRSSTEAN-ERECMGTCTGPPVPLMTGCE 352
301 PALSNDVADQDKQWVYVLAGCHVGYNDGHSRRESNARRECMCAVAPYPLMTGCE 360
301 PALSNDVADQDKQWVYVLAGCHVGYNDGHSRRESNARRECMCAVAPYPLMTGCE 360
353 AGFVNDGPPHTAFTPCGHVSESAKYSQIPLPHGTHAFHACPCATQOLVGEONCIK 412
353 AGFVNDGPPHTAFTPCGHVSESAKYSQIPLPHGTHAFHACPCATQOLVGEONCIK 412
361 PAFVVDGAFTHVAVPCGHVSEKSTRYMAEIPLPHGTHAFHACPCATQOLVLTQKAK 420
361 PAFVVDGAFTHVAVPCGHVSEKSTRYMAEIPLPHGTHAFHACPCATQOLVLTQKAK 420
413 LIFGSPID 420
413 LIFGSPID 420
421 LIFGSPID 428
421 LIFGSPID 428

RESULT 6
ID PELL1 HUMAN STANDARD; PRT; 418 AA.
AC Q96FA3; Q96SM0; Q96ZT5; Q9HCK0;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Pellino protein homolog 1 (pellino 1) (pellino related intracellular
DE signaling molecule).
GN Name=PELL1; Synonyms=PRISM;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Eumarchontoglires; Primates; Catarrhini; Homidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Thymus;
RX MEDLINE=21015032; PubMed=11132151; DOI=10.1007/s002510000249;
RA Rich T., Allen R.L., Lucas A.-M., Trowsdale J.,
RT "Pellino-related sequences from Caenorhabditis elegans and Homo
RT sapiens.";
Immunogenetics 52:145-149(2000).
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21203570; PubMed=11306823;
RA Resch K., Jockusch H., Schmitt-John T.,
RT "Assignment of homologous genes, PELL1/PELL1 and PELL2/PELL2, for the
RT pelle adaptor protein Pellino to mouse chromosomes 11 and 14 and human
RT chromosomes 2p3.3 and 14q21, respectively, by physical and radiation
RT hybrid mapping.";
Cytogenet. Cell Genet. 92:172-174(2001).
[3]
RP NUCLEOTIDE SEQUENCE.
RA Kennedy E.J., Moynagh P.N.,
RT "PRISM, a novel mediator of Toll/IL-1 signalling.";
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Ovary, Placenta, and Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schmech G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schmech A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 78-418.
RC TISSUE=Testis; carcinoma;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohyashi M., Nishi T., Shidohara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,

RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotta T.,
 RA Kusano J., Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Maehuo Y., Yamaeshita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [6]
 RP FUNCTION. AND INTERACTIONS WITH IRAK1, IRAK4 AND TRAF6.
 RX MEDLINE=2258429; PubMed=12496252; DOI=10.1074/jbc.M21212200;
 RA Jiang Z., Johnson H.J., Nie H., Qin J., Bird T.A., Li X.;
 RT "Pellino 1 is required for interleukin-1 (IL-1)-mediated signaling
 RT through its interaction with the IL-1 receptor-associated kinase 4
 RT (IRAK)-IRAK-tumor necrosis factor receptor-associated factor 6
 RT (TRAF6) complex.";
 RL J. Biol. Chem. 278:10952-10956(2003).
 RN [7]
 RP INTERACTIONS WITH TRAF6 AND MAP3K7.
 RX MEDLINE=22689054; PubMed=12804775; DOI=10.1016/S0014-5793(03)00533-7;
 RA Jensen L.B., Whitehead A.S.;
 RT "Pellino2 activates the mitogen activated protein kinase pathway.";
 RL FEBS Lett. 545:199-202(2003).
 CC -1- FUNCTION. Scaffold protein involved in the IL-1 signaling pathway
 CC via its interaction with the complex containing IRAK kinases and
 CC TRAF6. Required for NF-kappa-B activation and IL-8 gene expression
 CC in response to IL-1.
 CC -1- SUBUNIT: Found in a complex containing TRAF6, IRAK1 and IRAK4.
 CC -1- INTERACTS WITH MAP3K7.
 CC -1- SIMILARITY: Belongs to the pellino family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC -----
 CC EMBL: AJ278859; CAC04320.1; -; mRNA.
 CC EMBL: AF302505; AAG15393.1; -; mRNA.
 CC EMBL: AF300987; AAG17451.1; -; mRNA.
 CC EMBL: BC011419; AAH17419.1; ALT INIT; mRNA.
 CC EMBL: BC050019; AAH50019.1; -; mRNA.
 CC EMBL: BC063611; AAH63611.1; -; mRNA.
 CC EMBL: AK027668; BAB55280.1; ALT_INIT; mRNA.
 CC IntAct: Q96FA3; -;
 CC DR Ensembl: ENSG00000197329; Homo sapiens.
 CC DR HGNC: HGNC:8827; PELI1.
 CC DR InterPro: IPR006800; Pellino.
 CC DR PANTHER: PTHR12098; Pellino; 1.
 CC DR Pfam: Pf04710; Pellino; 1.
 CC FT CONFLICT 11 S -> F (in Ref. 1 and 5).
 CC FT CONFLICT 260 S -> P (in Ref. 5).
 CC SQ SEQUENCE 418 AA; 46286 MW; 233318A45E7546F7 CRC64;

QY 1 MESPGEERHAPNKEPVYKGLVLTGNGALPNGDRGRKRSFALYKRPKANGKSTYH 60
 DB 1 MFSPDQENH--PSKAPVKYKGLVLTGNGSLPNGDRGRKRSFALYKRPKANGKSTYH 58
 QY 61 VISTPQASKAISCKGHSISYTLSRNQTIVVEYTHDKQDMQVGVSTSPIDFVVTDI 120
 DB 59 IACTPQAAAIKSKDHSISYTLSRQIVVEYTHDSNTDMQVIGSTSPIDFVVTDI 118
 QY 121 SSGQNTDEAQITQSTSRACRIVCDRNPETARIPAAFGDSKNIPLFGEKAKWKNPDD 180
 DB 119 PSQSNSDPQSVOSTISRACRICRNPETARIPAAFGDSKNIPLFGEKAKWKNPDD 178
 QY 181 HMDGLTTNGVLVMPHPRGFTESQPCWREISVCGDVYTLRETRSAQGRKLVESSTNV 240
 DB 179 QNDGLTTNGVLVMPHPRGFTESQPCWREISVCGDVYTLRETRSAQGRKLVESSTNV 238
 QY 241 QDQSLDLCGATLMTAAGLFTTPQKILEALROGINARQCPGLNTLAPPSINRKE 300
 DB 239 QDQSLDLCGATLMTAAGLFTTPQKILEALROGINARQCPGLNTLAPPSINRKE 298
 QY 301 VVEKQPMAYLSCGHVGHYNNMGHRSDEANRECPMCRTPGVYPLMTGCEAGFYVDAG 360
 DB 299 VDEKQPMAYLSCGHVGHYNNMGHRSDEANRECPMCRTPGVYPLMTGCEAGFYVDAG 358
 QY 361 PPTAFTPCGHVCESEKSAKYWSQIPLPHGTTHAFAACPCATQVGEQNCIKLIFQGPID 420
 DB 359 PPTAFTPCGHVCESEKSAKYWSQIPLPHGTTHAFAACPCATQVGEQNCIKLIFQGPID 418

RESULT 7
 ID Q53T26 HUMAN PRELIMINARY; PRT; 418 AA.
 AC Q53T26;
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
 DE 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
 GN Hypothetical protein PELI1.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lesley K., Haekenson W., Martinka S.;
 RT "The sequence of Homo sapiens BAC clone RP11-547M24.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Waterston R.H.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Waterston R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RA Wilson R.K.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC012368; AA242474.1; -; Genomic DNA.
 KM Hypothetical protein.
 SQ SEQUENCE 418 AA; 46286 MW; 233318A45E7546F7 CRC64;

Query Match 83.7%; Score 1917; DB 2; Length 418;
 Best Local Similarity 81.4%; Pred. No. 2,1e-151;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

QY 1 MESPGEERHAPNKEPVYKGLVLTGNGALPNGDRGRKRSFALYKRPKANGKSTYH 60
 DB 1 MFSPDQENH--PSKAPVKYKGLVLTGNGSLPNGDRGRKRSFALYKRPKANGKSTYH 58
 QY 61 VISTPQASKAISCKGHSISYTLSRNQTIVVEYTHDKQDMQVGVSTSPIDFVVTDI 120
 DB 59 IACTPQAAAIKSKDHSISYTLSRQIVVEYTHDSNTDMQVIGSTSPIDFVVTDI 118
 QY 121 SSGQNTDEAQITQSTSRACRIVCDRNPETARIPAAFGDSKNIPLFGEKAKWKNPDD 180
 DB 119 PSQSNSDPQSVOSTISRACRICRNPETARIPAAFGDSKNIPLFGEKAKWKNPDD 178
 QY 181 HMDGLTTNGVLVMPHPRGFTESQPCWREISVCGDVYTLRETRSAQGRKLVESSTNV 240
 DB 179 QNDGLTTNGVLVMPHPRGFTESQPCWREISVCGDVYTLRETRSAQGRKLVESSTNV 238
 QY 241 QDQSLDLCGATLMTAAGLFTTPQKILEALROGINARQCPGLNTLAPPSINRKE 300
 DB 239 QDQSLDLCGATLMTAAGLFTTPQKILEALROGINARQCPGLNTLAPPSINRKE 298
 QY 301 VVEKQPMAYLSCGHVGHYNNMGHRSDEANRECPMCRTPGVYPLMTGCEAGFYVDAG 360
 DB 299 VDEKQPMAYLSCGHVGHYNNMGHRSDEANRECPMCRTPGVYPLMTGCEAGFYVDAG 358
 QY 361 PPTAFTPCGHVCESEKSAKYWSQIPLPHGTTHAFAACPCATQVGEQNCIKLIFQGPID 420
 DB 359 PPTAFTPCGHVCESEKSAKYWSQIPLPHGTTHAFAACPCATQVGEQNCIKLIFQGPID 418

61 VISTPQASKAISCKGHSISYTLSRNQTIVVEYTHDKQDMQVGVSTSPIDFVVTDI 120

Db 119 PGSQNSDTSQVSTISRACRIICERSPPFTARIYAAGDSKNTFLGKAAKMTSDG 178
 QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQGRKLVESNTVL 240
 Db 179 QMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQGRKLVESNTVL 238
 QY 241 QDGSLLDLCGATLLMRTADGLFTPTQKHEALROEINARPOCPVGLNTLAPPSINRKE 300
 Db 239 QDGSLLDLCGATLLMRTADGLFTPTQKHEALROEINARPOCPVGLNTLAPPSINRKE 298
 QY 301 VVEKOPMAYLSCGHVHGYNMGRSOTENRECEPCRCRVGVVPLMLGCEAGFYVDAG 360
 Db 299 VVDEKOPMAYLSCGHVHGYNMGRSOTENRECEPCRCRVGVVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHCSEKSAKYMSQIPLPHGTHAFHACPCFATQOLVGEONCIKILFQGPID 420
 Db 359 PPTHAFSPCGHCSEKSAKYMSQIPLPHGTHAFHACPCFANQOLAGEQYIRILFQGPID 418

RESULT 9
 Q5SRW7_MOUSE PRELIMINARY; PRT; 418 AA.
 AC O5SRW7_MOUSE
 DT 10-MAY-2005 (TREMBlrel. 30, Created)
 DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
 DE Pellino 1.
 GN Name=Pell1; ORFNames=RP23-93011.1-001;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 NCBI_TaxId=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lovell J.;
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL669979; CA126089.1; -; Genomic DNA.
 SQ SEQUENCE 418 AA; 46259 MW; 26072654577BBF7 CRC64;

Query Match 83.5%; Score 1912; DB 2; Length 418;
 Best Local Similarity 81.2%; Pred. No. 5.6e-151;
 Matches 341; Conservative 39; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSQGEBCAPKEPVKYGELVVLGYNALPNGDGRGRKSRPALYKRPKANGVKSTYH 60
 Db 1 MFSQGEBCAPKEPVKYGELVVLGYNALPNGDGRGRKSRPALYKRPKANGVKSTYH 58
 QY 61 VISTPQASKAISCKGHSISYTLISRNQTVVEYTHDKDTMFOVGRSTESPIDFVVYDTI 120
 Db 59 IACTPQAAKAIKNDQHSISYTLISRAQTVVEYTHDSNTDMFOIGRSTESPIDFVVYDTI 118
 QY 121 SSGQNTDEAQITOSTISRACRIICERNPFTARIYAAGDSKNTFLGKAAKMTSDG 180
 Db 119 PGSQNSDTSQVSTISRACRIICERSPPFTARIYAAGDSKNTFLGKAAKMTSDG 178
 QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQGRKLVESNTVL 240
 Db 179 QMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQGRKLVESNTVL 238
 QY 241 QDGSLLDLCGATLLMRTADGLFTPTQKHEALROEINARPOCPVGLNTLAPPSINRKE 300
 Db 239 QDGSLLDLCGATLLMRTADGLFTPTQKHEALROEINARPOCPVGLNTLAPPSINRKE 298
 QY 301 VVEKOPMAYLSCGHVHGYNMGRSOTENRECEPCRCRVGVVPLMLGCEAGFYVDAG 360
 Db 299 VVDEKOPMAYLSCGHVHGYNMGRSOTENRECEPCRCRVGVVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHCSEKSAKYMSQIPLPHGTHAFHACPCFATQOLVGEONCIKILFQGPID 420
 Db 359 PPTHAFSPCGHCSEKSAKYMSQIPLPHGTHAFHACPCFANQOLAGEQYIRILFQGPID 418

RESULT 10
 Q7ZXU3_XENLA PRELIMINARY; PRT; 418 AA.
 AC Q7ZXU3_XENLA
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Pell1-prov protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 NCBI_TaxId=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusik A., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uebli T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallory S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Rahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalka U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Merra W.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT Initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC041177; AAH44117.1; -; mRNA.
 DR InterPro; IPR006800; Pellino.
 DR Pfam; PF04710; Pellino; 1.
 SQ SEQUENCE 418 AA; 46231 MW; 04DCF12560BAD530 CRC64;

Query Match 82.3%; Score 1884; DB 2; Length 418;
 Best Local Similarity 80.5%; Pred. No. 1.2e-148;
 Matches 338; Conservative 39; Mismatches 41; Indels 2; Gaps 1;

QY 1 MFSQGEBCAPKEPVKYGELVVLGYNALPNGDGRGRKSRPALYKRPKANGVKSTYH 60
 Db 1 MFSQGEBCAPKEPVKYGELVVLGYNALPNGDGRGRKSRPALYKRPKANGVKSTYH 58
 QY 61 VISTPQASKAISCKGHSISYTLISRNQTVVEYTHDKDTMFOVGRSTESPIDFVVYDTI 120
 Db 59 IACTPQAAKAIKNDQHSISYTLISRAQTVVEYTHDSNTDMFOIGRSTESPIDFVVYDTI 118
 QY 121 SSGQNTDEAQITOSTISRACRIICERNPFTARIYAAGDSKNTFLGKAAKMTSDG 180
 Db 119 PGSQNSDTSQVSTISRACRIICERNPFTARIYAAGDSKNTFLGKAAKMTSDG 178
 QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQGRKLVESNTVL 240

```

Db 179 QMGSLTNGVLVHPRGFTEDSKGVMREISVCGNFSLRRETRSAQRKMWENESNEL 238
Qy 241 QDGSLLDLCGATILMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAPPSINRKE 300
Db 239 QDGSLLDLCGATILMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAPPSINRKE 298
Qy 301 VVEKOPMAVYLSGCHVGYNNMGHRSPTANERSCPCRTVGPVPLMTGCEAGFYVDAG 360
Db 299 VVDKOPWVYLVNGCHVGYNNMGHRSPTANERSCPCRTVGPVPLMTGCEAGFYVDAG 358
Qy 361 PPTHAFPCGHCSEKSAKTYWSOILPHGTHFAHAACPATOLVGEONCIKILFQSPID 420
Db 359 PPTHAFPCGHCSEKSAKTYWSOILPHGTHFAHAACPATOLVGEONCIKILFQSPID 418

RESULT 11
05ZKT7 CHICK PRELIMINARY; PRT; 405 AA.
AC 05ZKT7_
DT 25-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DE Hypochemical protein.
ORFNames=RCJMB04.944;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_TaxID=99883;
[1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicand S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Desilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Aichouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Castolico L., Poulin J., De Bernardis V.,
RA Cruaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kelis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crollius H.,
RT "genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.",
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014597; CAG00303.1; -; Genomic DNA.
SQ SEQUENCE 405 AA; 44790 MW; 8928FE5E52DB3BF CRC64;

Query Match 81.4%; Score 1864; DB 2; Length 405;
Best Local Similarity 81.8%; Pred. No. 5.5e-147;
Matches 333; Conservative 36; Mismatches 36; Indels 2; Gaps 1;

Qy 1 MESPQGEHCAPNKEPVYKGLVVLGYNGALPNDGRGRKRSRFPALYKRPKANGVPESTVH 60
Db 1 MESPQDENH--PSKAPVYKGLVVLGYNGSLPNDGRGRKRSRFPALYKRPKANGVPESTVH 58
Qy 61 VISTPQASKISCKGSHSYTLSSRNOTVVEYTHDQDTMFOVGRSTESPIDPVYTDI 120
Db 59 IACPQAKAISNDQSHSYTLSSRAQTVAVEYTHDNTMFOVGRSTESPIDPVYTDV 118
Qy 121 SSGSNTDEAOITOSTISRFACRIYCDNREPYTAIFAGFDSKNIFLGEKAAKMPNDG 180
Db 119 PGSQNSNDTOSVSTISRFACRIICENRPFTARIYAAGFDSKNIFLGEKAAKMPNDG 178
Qy 181 HMDGLTNGVLVHPRGFTEDSKGVMREISVCGDVTTLRETRSAQRKMWENESNEL 240
Db 179 QMGSLTNGVLVHPRGFTEDSKGVMREISVCGDVTTLRETRSAQRKMWENESNEL 238
Qy 241 QDGSLLDLCGATILMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAPPSINRKE 300
Db 239 QDGSLLDLCGATILMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAPPSINRKE 298
Qy 301 VVEKOPMAVYLSGCHVGYNNMGHRSPTANERSCPCRTVGPVPLMTGCEAGFYVDAG 360
Db 299 VVDKOPWVYLVNGCHVGYNNMGHRSPTANERSCPCRTVGPVPLMTGCEAGFYVDAG 358

```

```

Qy 361 PPTHAFPCGHCSEKSAKTYWSOILPHGTHFAHAACPATOLVGE 407
Db 359 PPTHAFPCGHCSEKSAKTYWSOILPHGTHFAHAACPATOLVGE 405

RESULT 12
04SGC9_TETNG PRELIMINARY; PRT; 428 AA.
AC 04SGC9_
DT 13-SEP-2005 (TEMBLrel. 31, Created)
DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DE Chromosome 17 SCARF4597, whole genome shotgun sequence.
ORFNames=GSTENG00018701001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_TaxID=99883;
[1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicand S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Desilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Aichouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Castolico L., Poulin J., De Bernardis V.,
RA Cruaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kelis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crollius H.,
RT "genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.",
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014597; CAG00303.1; -; Genomic DNA.
SQ SEQUENCE 428 AA; 47639 MW; E4187096D4454577 CRC64;

Query Match 78.1%; Score 1788; DB 2; Length 428;
Best Local Similarity 76.9%; Pred. No. 1.3e-140;
Matches 329; Conservative 38; Mismatches 53; Indels 8; Gaps 3;

Qy 1 MESPQGE--EHCAPNKEPVYKGLVVLGYNGALPNDGRGRKRSRFPALYKRPKANGVPESTVH 58
Db 1 MESPQDENH--PSKAPVYKGLVVLGYNGSLPNDGRGRKRSRFPALYKRPKANGVPESTVH 56
Qy 59 VHIISTPQASKISCKGSHSYTLSSRNOTVVEYTHDQDTMFOVGRSTESPIDPVYTDI 118
Db 61 VHAACSPQAKAISNDQSHSYTLSSRAQTVAVEYTHDNTMFOVGRSTESPIDPVYTDI 120
Qy 119 TISGNTDEAOITOSTISRFACRIYCDNREPYTAIFAGFDSKNIFLGEKAAKMPNDG 176
Db 121 TVAGSSNDTOSVSTISRFACRIICENRPFTARIYAAGFDSKNIFLGEKAAKMPNDG 180
Qy 177 NPGGNDGLTNGVLVHPRGFTEDSKGVMREISVCGDVTTLRETRSAQRKMWENESNEL 236
Db 181 TFGQNDGLTNGVLVHPRGFTEDSKGVMREISVCGDVTTLRETRSAQRKMWENESNEL 234
Qy 237 TVNLQDGSLLDLCGATILMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAPPSINRKE 296
Db 241 SHELVQDGSLLDLCGATILMTADGLFHTPTQKHIALROEINARPOCPVGLNTLAPPSINRKE 294
Qy 297 NRKQVVEKOPMAVYLSGCHVGYNNMGHRSPTANERSCPCRTVGPVPLMTGCEAGFYVDAG 352

```

DB 301 HRKDPEKOPWYLLQGHVGHYHNMGNPREBERGREGDRCEPMCRKTGPPVPLMLGCE 360

QY 353 AGFYVDAGPPTHAFTPCGHVCESEKSAKCYMSOILPLPHGTHFAHACPCATOLVGEONCIK 412

DB 361 AGFYVDAGPPTHAFTPCGHVCESEKSAKCYMSOILPLPHGTHFAHACPCATOLVGEONCIK 420

QY 413 LIFOGPID 420

DB 421 LIFOGPID 428

RESULT 13

Q4SBZ1.TETNG PRELIMINARY; PRT; 440 AA.

ID Q4SBZ1.TETNG

AC Q4SBZ1.TETNG

DT 13-SEP-2005 (T-EMBLrel. 31, Created)

DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)

DE Chromosome 14 SCAF14660, whole genome shotgun sequence.

DE (Fragment).

GN ORFNames=GSTENG0020745001;

OS Tetradon nigrovittidis (Green puffer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

OC Tetraodontidae; Tetraodontidae; Tetradon.

NCBI_TaxID=99893;

[1]

RP NUCLEOTIDE SEQUENCE.

RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,

RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,

RA Micaud S., Jaffe D., Fisher S., Lutfalla G., Dosset C., Segurens B.,

RA Desilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,

RA Anthouard V., Jubin C., Castellid V., Karinka M., Vacherie B.,

RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,

RA Cnuand C., Duprat S., Brottier P., Couanceau J.P., Gouzy J.,

RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,

RA Kellis M., Wolff J.W., Guiso R., Zody M.C., Mesirov J.,

RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,

RA Lauder V., Schacher V., Quetier F., Saurin W., Scapellato C.,

RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,

RT Genome duplication in the teleost fish Tetradon nigrovittidis reveals

RL the early vertebrate proto-karyotype.";

RT Nature 431:946-957(2004).

RA 12

RP NUCLEOTIDE SEQUENCE.

RG Genoscope, Whitehead Institute Centre for Genome Research;

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

CC -1- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; CA01014660; CAC01841.1; -; Genomic_DNA.

FT NON_TER 1

SQ SEQUENCE 440 AA; 4845 MW; A2DB6C7CF9E85099 CRC64;

Query Match 71.9%; Score 1647; DB 2; Length 440;

Best Local Similarity 67.9%; Pred. No. 8.4e-129;

Matches 299; Conservative 47; Mismatches 62; Indels 32; Gaps 4;

QY 12 PNKEPVKYGELVVLGNGALPNDRGRKSRFALYKPKANGVPSFTVHVISFPAQAKAI 71

DB 4 PVMDPIYKGLVILVINGSLPSGDRGRKSRFALYKPKANGVPSFTVHVISFPAQAKAI 63

QY 72 SCGQHSISTYLRNQNQVVEYTHDKTDMFQVGRSTESPIDFVYDTTSGSQNTDAQI 131

DB 64 HSSGQHSISTYLRNQNQVVEYTHDKTDMFQVGRSTESPIDFVYDTTSGSQNTDAQI 123

QY 132 TOSTISFAPRIYCDREPEYTAIFAAGPSSKNIIFGEAKAKKNDGADGLTTNGVL 191

DB 124 APSTISFACRIYCDREPEYTAIFAAGPSSKNIIFGEAKAKKNDGADGLTTNGVL 183

QY 192 VMHPRGGFTESQPGVWRISVCGDVYTLRETRSAOQRGLV----- 233

DB 184 VMHPRGGFTESQPGVWRISVCGDVYTLRETRSAOQRGLV----- 242

QY 234 ---ESEFNVLQDGLSLDLGATLMTADGLFTTPQKHIEALROEINARPOCPVLNT 290

DB 243 LQAEGETSALRDSGLVDLCATLMTAGBLNAPLTHLEALRQELNLSRPPCVGLNT 302

QY 291 LAFPSINRKEVNEKOPWYLLQGHVGHYHNMGNPREBERGREGDRCEPMCRKTGPPVPLMLGCE 341

DB 303 LAFPSINRKEVNEKOPWYLLQGHVGHYHNMGNPREBERGREGDRCEPMCRKTGPPVPLMLGCE 362

QY 342 GPVYPLMLGCEPAPVYDAGPPTHAFTPCGHVCESEKSAKCYMSOILPLPHGTHFAHACPCFA 401

DB 363 GPVYPLMLGCEPAPVYDAGPPTHAFTPCGHVCESEKSAKCYMSOILPLPHGTHFAHACPCFA 422

QY 402 TOLVGEONCIKLFQGPID 420

DB 423 AAL-GSPGWTLLIFQGPID 440

RESULT 14

P6133_MOUSE STANDARD; PRT; 445 AA.

ID P6133_MOUSE

AC P6133_MOUSE

DT 29-MAR-2004 (Rel. 43, Created)

DT 29-MAR-2004 (Rel. 43, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Pellino protein homolog 3 (Pellino 3).

DE Name=Pellino3;

GN Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

[1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC STRAIN=CS7BL/6J; TISSUE=Retina;

RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;

RA Okazaki I., Furuno M., Kasukawa T., Adachi Y., Bono H., Kondo S.,

RA Nikaide I., Furuno M., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,

RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schirml L.M., Kanapin A., Matsuda H., Batelou S., Beisel K.W.,

RA Blake J.A., Bird D., Brusic V., Choithia C., Corbani L.E., Cousins S.,

RA Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Guenichon S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,

RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,

RA Maglocz D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Numata K., Okido T., Pavan W.J., Pereira G., Pesole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Sempie C.A., Seton M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,

RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Wangabe Y., Wells C.,

RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang L., Yang L.,

RA Yuan Z., Zaytoun M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,

RA Hirokane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,

RA Hara A., Haseizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,

RA Yeanishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.,

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

[2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC TISSUE=Mammary tumor;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Straussberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,

[illegible]

30	88.5	3.9	110.4	2	138869	transcription factor
31	88.5	3.9	112.1	2	S54504	hypothetical protee
32	88	3.8	81.2	2	AC23439	hypothetical prote
33	88	3.8	134.5	2	H90975	hypothetical prote
34	87.5	3.8	54.0	2	J34699	cadmium, zinc, cob
35	87.5	3.8	72.4	2	A48569	antigen Em100 - B1
36	87.5	3.8	165.0	2	S28721	hypothetical prote
37	87	3.8	31.5	2	T03681	putrescine N-methyl
38	87	3.8	92.3	2	G81253	probable oxidoredu
39	86.5	3.8	67.1	2	J80288	kruppel-type zinc
40	86	3.8	51.8	1	N18CBT	nitrrogenase (BC 1,
41	86	3.8	61.0	2	T22687	hypothetical prote
42	86	3.8	84.8	2	C84845	probable salt-indu
43	86	3.8	133.6	2	B83770	hypothetical prote
44	86	3.8	155.5	2	S46224	peroxidasin - frui
45	86	3.8	174.4	1	S18644	multifunctional am

ALIGNMENTS

RESULT 1

hypothetical protein F25B4.2 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Dec-2004
 C:Accession: T25731
 R:Gatung, S.
 submitted to the EMBL Data Library, July 1996
 A:Description: The sequence of C. elegans cosmid F25B4.
 A:Reference number: Z25076
 A:Accession: T25731
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-484 <GAT>
 A:Cross-references: UNIPROT:Q22967; UNIPARC:UPI0000179271; EMBL:U64842; PIDN:AA837081.1,
 A:Experimental source: strain Bristol N2; clone F25B4
 C:Genetics:
 A:Gene: CESP:F25B4.2
 A:Map position: 5
 A:Introns: 35/2, 78/3, 112/3, 136/1, 191/3, 334/3, 406/2, 437/3

Query Match: 38.9%; Score 890; DB 2; Length 484;

Best Local Similarity 43.5%; Pred. No. 1.7e-64;

Matches 178; Conservative 68; Mismatches 143; Indels 20; Gaps 7.

10. КОПИИ СЪСТАВЛЯЮТ ЧАСТЬ ДОКУМЕНТА. КОПИИ НЕ ДОЛЖНЫ БЫТЬ ПОДАВАЕНЫ НА СТОЛОВАХ ИЛИ НА ПУНКТОВЫХ СТОЛОВАХ.

13 IGELV LG INGHLE FN GDRKKA - NOKF MW I NKF MW G V N F S I V H V S I F Y M D M A D C M O G H I I

Db . 28 YGELLGNGAENRATSKRYLTKKVLRRRDSANGIKKCTVHNVSTSD-TKLTKDKARH 86

78 SISYTLNRQTVVEYTHDKDIDMFQVGRSTESPIDFVITD-1SGSQNTDEA-----QI 131

[illegible]

D6 87 TVSFHSDSNKSVIEYADPSKMFQIGRASDDQIDFTVIDTMMFLPEHSDA VPARPQI 146

133 TOSTERACRYL/CDBNBPYTABTBAAGPNSKNTIIGKAKTKNPNCHMDG 184

[illegible]

Db 147 DYLEKGDRTSTISRFACRIITDRENSNKAYLYAAGFDAHONISINKKSLKWTKSNGEVDG 206

185 LITNGVLVHPRG---GFTESQPGVWREISVCGDYTLRETRSAQQRGKLVESETNL 240

[illegible]

Db 207 LITNGVLLHPNKDLDLDIVDKPMYKKWREVSINGDVYEPRVTRSSSAKGVFVPEWIML 266

0- 341 00687 INI CSATET IWBETACI ENEPTOKUTENI BOITNABPOCBYCI NTI AEPSTNPKE 300

241 QDGSILDECGAILMKIADGDFHIFIVNIEEUNGEINAKKFCFVQUNIMAFESINAKS JVV

267 ODGTTIDLCATTIWRATDGLERSPKMRELEMALDRIINAGRPCPVNTLTVIPKKRNGR 326

1. **Introduction**

301 VVEEKOPWYLSGCHVHGYNWGHRSDEANERECPCMCRTVGP-YVPLWLGCBAQFYVDA 359

[illegible]

Db 327 QINRRQPYVYLQCGHVGRHEWGVQENSGQRSGKCPICLVESERIVQLSMGMEPSFHLDS 386

400 פסוקי תורה ונביאים

360 GPFTHAFIPCGHVCSEKSAKYSQLEPHGTHAFHACPFCAIQLVGEQ 408

387 GVT.DHTFNPBGHMA SKOTVT.YWSP.TPI.POGTGPYDVPVCP ECVOI.TATER 435

RESULT 2

T28405
ORF MSV244 hypothetical protein - Melanoplus sanguinipes entomopoxvirus
C/Species: Melanoplus sanguinipes entomopoxvirus
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C/Accession: T28405
R/Alonso, C.L.; Tulman, E.R.; Lu, Z.; Oms, E.; Kuteh, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A/Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A/Reference number: Z20484; MUID:99102612; PMID:9847359
A/Accession: T28405
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-215 <ARO>
A/Cross-references: UNIPROT:Q9YVJ8, UNIPARC:UPI00000F81BF, EMBL:AF063866, NID:g4049647,
C/Genetics:
A/Note: MSV244

Query Match 5.3%; Score 122.5; DB 2; Length 215;
Best Local Similarity 18.7%; Pred. No. 0.016;
Matches 61; Conservative 31; Mismatches 84; Indels 151; Gaps 13;

QY 81 YTLRNQTVVVEYTHDKDTMFQVGRSTESPIDFVVTDTISGSQNTDEAOITOSTISRA 140
DB 19 YNLIYNNK-----NYT-----FTIGRN--SKCDIFIND-----KKISNIA 50
QY 141 CRIVCDNRPETATIFAPGPDSSKNITLGEKAKAKNPDGMGLTTNGVLVNHPRGGFT 200
DB 51 CIHKCD-YEKKCKPIYGGCSNKKYLYPDNN-----DGLQNGIFIKFPN----- 94
QY 201 EESQPGVWRREISVCGDYVTLRETSAAQQRGLVSESTNVLDGSLIDLCGATLLMRTADG 260
DB 95 -----LEMDLSTINGNITYL-----KSNIKKLINELIDTLIDISGNIYWRSI-- 139
QY 261 LFTPTQKHTEALROEINAAAPQCPVGLNTLAFPSINRKEVEBKOPWAYLSCGHVGYH 320
DB 140 ----- 139
QY 321 NWGHRSTEANEREC--PMCRTPGVYPLWLCGAGTYVDAGPPTAFTPGHVCSEKSA 378
DB 140 -----DVKLNKKICNYSMC-TSTKY-----EYMFYKCGHKINKKTA 174
QY 379 KYWSQI-----PLPHGTHAFHAACPCF 400
DB 175 TFWKKTMLCNDIYTKYKILNICPCF 201

RESULT 3

A54146
invasion-inducing protein Tiam-1 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C/Accession: A54146
R/Habers, G.G.M.; Scholtes, E.H.M.; Zuydgeest, D.; van der Kammen, R.A.; Stam, J.C.; Bet
Cell 77, 537-549, 1994
A/Title: Identification of an invasion-inducing gene, Tiam-1, that encodes a protein wit
A/Reference number: A54146; MUID:94243921; PMID:7999144
A/Accession: A54146
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1591 <HAB>
A/Cross-references: UNIPROT:Q60610, UNIPARC:UPI0000028188, GB:U05245; NID:g497638; PIDN:
F/1040-1234/Domain: CDC24 homology <CD24>

Query Match 4.5%; Score 103.5; DB 2; Length 1591;
Best Local Similarity 22.1%; Pred. No. 8.7;
Matches 81; Conservative 53; Mismatches 166; Indels 67; Gaps 18;

QY 3 SPGEHECAPNKEVKYGEVLVINGALPNDGRRKSRPALYKRPKANGVKPSTVAVI 62
DB 219 SPQLSTC---GRANSGLDLYAQKNSGVKANG---PRNRPSSYCRMIVSDIPDLAKHKM 272

QY 63 -----STPOAS--KAISCGHISISTYLSRNQTV---VEYTHDKDTMFQVGRSTRE-- 109
DB 273 PPAABETTPPYSTNTYTLPCRKHCHLSEGATNPOLISKSKQGRRAKTTQDVNTGESSEFA 332
QY 110 -SPIDFVVTDTISGSQNTDEAOITOSTISRFACRIVCDNRPETATIFAPGPDSSKNITFL 168
DB 333 DSGIEGATTTDLSRSN--ATNSYS-----PTGNAF-VGSDSGSS-SF 375
QY 169 GEKAAK--KKNPGHMDGLTTNGVLVNHPRGGTTE-----SQGVWRREISVCGDYVTL 220
DB 376 GDRARQGVENFRRELEMSSTYNSLEEGASASDQSGTSLSPG-QSDILLTLTAQGV 434
QY 221 RETRS-----AQQRGLVSESTNVLDGSLIDLCGATLLMRTAD---GLFPTPTQKH 270
DB 435 KKGALAAVKNFLVHKQKVKVESATRRKMKHYVSLGCTLFFYETGRSGSIDNSVPKHA 494
QY 271 EALROBINARPOCPVG-----LNTLAFPSINRKEVEBKOPW--AYLSCGHVGYH 321
DB 495 VVWENSIVQAVPEHPKDDFVFCLSNLSGDAFLPQTTSQTELEBWITAHISACAAVARRH 554
QY 322 WGRSDT 328
DB 555 --HKEDT 559

RESULT 4

T03447
dihydrokaempferol 4-reductase (EC 1.1.1.219) A - sorghum
N/Alternate names: dihydroflavonol 4-reductase; NADPH-dependent reductase A1-a
C/Species: Sorghum bicolor (sorghum)
C/Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004
C/Accession: T03447
R/Chen, M.; Samiguel, P.; Bennetzen, J.L.
Genetics 148, 435-443, 1998
A/Title: Sequence organization and conservation of Sh2/A1-homologous regions of sorghum
A/Reference number: Z14952; MUID:98133900; PMID:9475753
A/Accession: T03447
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-350 <CHE>
A/Cross-references: UNIPROT:P93776; UNIPARC:UPI00000A710A; EMBL:AF010283; NID:g2735839; I
C/Genetics:
A/Intons: 52/1; 108/3; 173/3
C/Superfamily: dihydrokaempferol 4-reductase
C/Keywords: oxidoreductase

Query Match 4.4%; Score 100; DB 1; Length 350;
Best Local Similarity 24.1%; Pred. No. 2.1;
Matches 47; Conservative 21; Mismatches 71; Indels 56; Gaps 9;

QY 168 LGKAAKMKNPDDGMDGLTTNGVLVNHPRGGFTBESQGVWRREISVCGDYVTLRETRSAQ 227
DB 1 MGEVVAITWETEG---GAGYKGPVVVTGASGFL-----GSLVMKLLQAGYTVRAT----- 48
QY 228 QRKGLVSESTNVLDGSLIDLCAT---LJL-----RTADGLFHTPTQKH 270
DB 49 -----VRDPANVATKTYPLDLPGATERLSLWKADLADBSFDDAATCGGVFAVATPMDF 103
QY 271 EALROBINARP-----QCPVG--LNTLAFPSINRKEVEBKOPWAYLSCGHVGYH 318
DB 104 ESKDPENEVYKPIYVEGMSIMRACKKAGYRTRVFTSSAGTVIERQ-----RPVYD 156
QY 319 YNHWHRSDTEANER 333
DB 157 QDNW---SDVDFCOR 168

RESULT 5

AI0072
probable autotransporter protein YPO0587 [imported] - Yersinia pestis (strain CO92)
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C/Accession: A10072
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Terraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dave, P.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
 Nature 413, 523-527, 2001
 A/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A/Reference number: AB0001; MUID:21470413; PMID:11586360
 A/Accession: A10072
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-994 <R>
 A/Cross-references: UNIPROT:Q9P286; UNIPARC:UPI00000DCE70; GB:AL590842; PIND:CA089444.1;
 C/Genetics:
 A:Gene: YPO0587

Query Match 4.3%; Score 98; DB 2; Length 994;
 Best Local Similarity 22.1%; Pred. No. 13;
 Matches 70; Conservative 53; Mismatches 139; Indels 54; Gaps 17;

QY 13 NKEPVK---YGLVVGYNALPNDG-RGRKSRPLKPKA-NGVKSTVAVISIPA 67
 DB 27 NKEPIATCAAAAFIFESSSVIANPDHGIIVGKSLINKQSAVNALINEGSLVLTDSA 86
 QY 68 SKAISCQGSISYTLISRNQTVVYTHDQDMFQVGRSTESPIDVVTDTISGS--QN 125
 DB 87 SAHTVAVNGSI-FTLKEDSTADIT---SYTGFSSLSGSSKANINTV---LSGGLV 138
 QY 126 TDEAQITOSTISRPACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKKNPDGND-- 183
 DB 139 NDDASTETITIS-----SDIEKSTVRLYDQ--SATVTYGDNDILVSGDSRAEET 189
 QY 184 GLTNGVLYVWHPRGPFEEBQPGVWEISVCGDYTLIRETRSAAQOKLYSESTNVLQDG 243
 DB 190 HVTGKGGLIYV-----SEQCGPLTKNTQTAGTL-TLKSVDLTGKTEFVSAT-IKTTG 241
 QY 244 SLIDLCATLMTADGLPHTPQKITEALRQEIINARPOCPGLMTLAPSRKVEVE 303
 DB 242 HLIDNQCQL-----IFNSDKDIYEAMIDGQSLTKENP-LTTLTSSAG----- 285
 QY 304 EKQPM--AVLSCGHV 317
 DB 286 --DAWVASVYSGETH 299

RESULT 6

C86547

polymorphic outer membrane protein H family [imported] - Chlamydia pneumoniae (strain
 C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C/Accession: C86547
 R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Iw
 Nucleic Acids Res. 28, 2311-2314, 2000
 A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A/Reference number: A86491; MUID:20330349; PMID:10871362
 A/Accession: C86547
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-978 <STO>
 A/Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI0000131C83; GB:BA000008; NID:g8978825; PI
 A/Experimental source: strain J138
 C/Genetics:
 A:Gene: pmp_14
 C/Superfamily: Polymorphic membrane protein H family

Query Match 4.2%; Score 95.5; DB 2; Length 978;
 Best Local Similarity 22.1%; Pred. No. 20;
 Matches 47; Conservative 25; Mismatches 74; Indels 67; Gaps 9;

QY 102 FQVGRSTESPIDVVTDTISGONTDEAQITOSTISRPACRIVCDRNE----- 149
 DB 523 FTIGKLAFPDFSLKRPVSAVAGTKNTY-----LTGALVLDHEDVDLYDMVSLQS 576
 QY 150 PYTARIFAAGFDSKNIFLGEKAAKKNPDGNDGLTTNG-----VLVWHPRG 197

DB 577 PVAIP-----AVFGATVTKGFPDGEIATRSHTGYGQKMSYTSRPLLIAPDG 627
 QY 198 GFPEESQP-----GWREISVCGDYTLIRETRSAAQOKLYSE-----TNVL 240
 DB 628 GFPGGSPSANTLYAVWNSDYLVRSTYTL-----DPERYGEIYVNSLWISFLGNQAFSDIL 683
 QY 241 QDGSLLIDLCATLMTADGL-----PHTPQKH 269
 DB 684 QDVLIDHPELSI-----TAKALGAVVHTPRQGH 713

RESULT 7

B81593

polymorphic membrane protein H family CP0298 [imported] - Chlamydia pneumoniae (stra
 C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C/Accession: B81593
 R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, I
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A/Title: Genome sequences of Chlamydia trachomatis McpN and Chlamydia pneumoniae AR39.
 A/Reference number: A81500; MUID:20150255; PMID:10684935
 A/Accession: B81593
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-978 <REA>
 A/Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI0000131C83; GB:AE002191; GB:AE002161; NID
 A/Experimental source: strain AR39, HL cells
 C/Genetics:
 A:Gene: CP0298
 C/Superfamily: Polymorphic membrane protein H family

Query Match 4.2%; Score 95.5; DB 2; Length 978;
 Best Local Similarity 22.1%; Pred. No. 20;
 Matches 47; Conservative 25; Mismatches 74; Indels 67; Gaps 9;

QY 102 FQVGRSTESPIDVVTDTISGONTDEAQITOSTISRPACRIVCDRNE----- 149
 DB 523 FTIGKLAFPDFSLKRPVSAVAGTKNTY-----LTGALVLDHEDVDLYDMVSLQS 576
 QY 150 PYTARIFAAGFDSKNIFLGEKAAKKNPDGNDGLTTNG-----VLVWHPRG 197
 DB 577 PVAIP-----AVFGATVTKGFPDGEIATRSHTGYGQKMSYTSRPLLIAPDG 627
 QY 198 GFPEESQP-----GWREISVCGDYTLIRETRSAAQOKLYSE-----TNVL 240
 DB 628 GFPGGSPSANTLYAVWNSDYLVRSTYTL-----DPERYGEIYVNSLWISFLGNQAFSDIL 683
 QY 241 QDGSLLIDLCATLMTADGL-----PHTPQKH 269
 DB 684 QDVLIDHPELSI-----TAKALGAVVHTPRQGH 713

RESULT 8

G72076

polymorphic outer membrane protein h family - Chlamydia pneumoniae (strain CML029)
 C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C/Accession: G72076
 R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Greenwood, J.;
 Nature Genet. 21, 385-389, 1999
 A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A/Reference number: A72000; MUID:99206606; PMID:10192388
 A/Accession: G72076
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-978 <ARN>
 A/Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI000046AB; GB:AE001629; GB:AE001363; NID:
 A/Experimental source: strain CML029
 C/Genetics:
 A:Gene: pmp_14
 C/Superfamily: Polymorphic membrane protein H family

Query Match 4.2%; Score 95.5; DB 2; Length 978;
 Best Local Similarity 22.1%; Pred. No. 20;
 Matches 47; Conservative 25; Mismatches 74; Indels 67; Gaps 9;

102 FQVGRSTESPIDVVTDTISGSONTDEAQTOSTISRFACRYCDRNE-----149
 523 FTIGKIAFDPSFLKEDFVSASVNACTKNT-----LTGALVLDHEDVTDLYDMVSLQS 576
 150 PYTARIFPAGFSDSKNIFLGEKAAKMPDGHMDGLTNG-----VLVNHPRG 197
 577 PVAIPF-----AVFKATATVTKGFPDGEIATPSHHYQCKMSYMSRPLLIAPADG 627
 198 GFTEESOP-----GVWRISVCGDVYTLRETSAAOQKGLVSEB-----TNVL 240
 628 GFEGGSPSANTLYAVWNSSTLVRSYTL-----DPERYGEIVNSLMISFLGNQARSDDL 683
 241 QDGSLLIDLCGATLLMTADGL-----FHTPTQKH 269
 684 QDVLLIDHPGLSI---TAKALGAVERTPRQGH 713

RESULT 9

C81804
 alanine-crRNA ligase (EC 6.1.1.7) NMA1788 [Imported] - Neisseria meningitidis (strain Z24)
 C/Species: Neisseria meningitidis
 C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
 C/Accession: C81804
 R/Parikh, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 R.; Holroyd, S.; Uggelø, K.; Leather, S.; Moutle, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A/Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
 A/Reference number: A81775; MUID:20222556; PMID:10761919
 A/Accession: C81804
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-874 <PAR>
 A/Cross-references: UNIPROT:Q9JTG4; UNIPARC:UPI0000136315; GB:AL162757; GB:AL157959; NID
 A/Experimental source: serogroup A, strain Z2491
 C/Genetics:
 A/Gene: alas; NMA1788
 C/Superfamily: alanyl-crRNA ligase
 C/Keywords: ligase

Query Match 4.0%; Score 92.5; DB 2; Length 874;

Best Local Similarity 23.0%; Pred. No. 30; Mismatches 127; Indels 77; Gaps 15;

66 QASKAISCKQHSISYTLNRQT---VVVEYTHDKOTDMFQVGRSTESPIDFVVTDTISG 122
 441 KANAQLEPYDQDTEFKYSERQTESKYLALYKQGEVNLNEDSGAVVIDFPTFAESG 500
 123 SONTDEAQTOSTISRFACRYCDRNEPYTARIFA- GFDSSKNIFGE-----K 171
 501 GQVGDVGYIFSSR-NRFEVR---DTQKIRAAVFGQFQVTSGLKAGDSVTAKVDDIR 555
 172 AAKMKNDP-----GHMDGLTNGVVL-----MHRGFTSE-----202
 556 NANNRNNSATLHMKALRDVAGHVE--QKSLVTAESTRFPDISHPQAVTAEBIAEVER 612
 203 -SOPGVAREISVCGDVYTLRETSAAOQKGLV-----ESETVVLQDGSV-IDLCGATLL 254
 613 RVNEAVLVANVAVNAALISMED---AQKTGAMMLEGKYGGEVRLQWCGFSTELCGGTHV 669
 255 WRPAD-GLFHTPTQKHIEALROBINARPOCPVGLNTLAPPSINRKEVE-----E 304
 670 SRFGIDGLFRTISEGCIAGVRRIEAF---TGLNALKWAQEQERLVKDIIAETKAQTE 724
 305 KQPMAYLVSCGHV 317
 725 KDVLAKIQAGAAH 737

RESULT 10

S42373
 hypothetical protein T20G5.3 - *Caenorhabditis elegans*
 C/Species: *Caenorhabditis elegans*
 C/Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 09-May-2004
 C/Accession: S42373

R/Smith, A.
 submitted to the EMBL Data Library, March 1994
 A/Reference number: S42368
 A/Accession: S42373
 A/Molecule type: DNA
 A/Residues: 1-3051 <SMI>
 A/Cross-references: UNIPARC:UPI000017B8C8; EMBL:Z30423; NID:G458479; PID:G458485
 C/Genetics:
 A/Intons: 96/1; 166/3; 201/2; 355/1; 404/1; 427/2; 794/1; 892/1; 952/2; 1041/1; 1146/1;
 F/512-679/Domain: von Willebrand factor type A repeat homology <WAL>
 F/754-793/Domain: fibronectin type II repeat homology <2FI>
 F/1201-1244/Domain: EGF homology <EGF>

Query Match 4.0%; Score 92; DB 2; Length 3051;

Best Local Similarity 21.6%; Pred. No. 1,8e+02;
 Matches 94; Conservative 44; Mismatches 156; Indels 142; Gaps 24;

4 PQG---EEHCAPNK-----EP---VKYGEVLVLYNGALPNQDGRGRKRSFALYKR 48
 1687 PGRVCSAPREVCPENHDCSSAAVCEPLGGMKYQCVCIQGVYDOSPGSQKGRVCVRNACHD 1746
 49 PRKNGVYKPTVAVYISTPQASKAISCKQHSISYTLNRQTVVEYTH-DKQTDMPQVGRS 107
 1747 PRLN-----TC-SRNAICTDEPRGYRCECKGFMDRSPSSQGRV 1786
 108 TESP-----IDFVVTDTISGSONTDEAQTOSTISRFACRYCDRNE 149
 1787 CEPPPPPPPPRRHPCQDPERNDCHPACTGATAGQS-----YTCBCLSGVADRSPPRRK 1841
 150 PYTARIFA--AGFDSSKNIFLGEKAAKMPDGHMDGLTT--NG--VLVNHPRGFTES 203
 1842 PGRLLCYLTGPVCLDPDQON-----DCHAAALICSEVNGPEKYCKCKRDGYTDS 1888
 204 QPQVMEISVCGDVYTLRETSAAOQKGLVSESTNVLQDGSLLIDC-----CATLLMRA 258
 1889 PDLRRPRGRI-----KGLNCELDRLNCHSLAVCKDLPNGT-----1928
 259 DGLFHTPTQKHIEALROBINARPO--CPVGLNTLAFPSINR-----KEVVEKQPMAYLS 312
 1929 -----CCPTIAKQSPDPKPKGRICSLSVNECANPSLNSASAPADCFDENGY---R 1978
 313 CGHVGTH-----NMGHRSDEANE-----RECEMCTVGPVPLMLGCEAGFYV 357
 1979 CRCRNGYHDDDPAHPPGRCSFMINECDSSNLNDCBNANCIDTAGGYD---CACKAP-YR 2034
 358 DAGPPTHAFTPGCHVC 373
 2035 DEGPPO---SP-GRIC 2046

RESULT 11

T40476
 hypothetical protein SPBC4B4.04 - fission yeast (*Schizosaccharomyces pombe*)
 C/Species: *Schizosaccharomyces pombe*
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C/Accession: T40476

R/Beck, A.; Reinhardt, R.; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, May 1997
 A/Reference number: Z21932
 A/Accession: T40476
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-576 <BEC>
 A/Cross-references: UNIPROT:O74965; UNIPARC:UPI000006A650; EMBL:AL023706; PIDD:CAA19284.1
 A/Experimental source: strain 972h-; cosmid c4B4
 C/Genetics:
 A/Gene: SPDB:SPBC4B4.04

A:Map position: 2
A:Insertions: 9/1
C:Superfamily: translation initiation factor eIF-2A

Query Match 4.0%; Score 91.5; DB 2; Length 576;
Best Local Similarity 20.6%; Pred. No. 20;
Matches 56; Conservative 35; Mismatches 98; Indels 83; Gaps 12;

39 RKSRLALYKRPKNGV-----KPTVAVISPPQASKAISCGCHSISYTLRNQTV 90
3 OKQFA-YRSSKSGVLAENASVSPKFAISSPANNACVSPNG----- 47
91 VERTHDQTMFQGRSTESPIDVVTDTISGQNTDEAQTOSTISRPACRIYCDRNEP 150
48 -----LFAVATATQ-----VINDTSSG-----ALITQLP----- 72
151 YTAIRFAPGDSKNIFLGEKAKAKNPDGMDGLTTNGVLVMPHPRG-----FTESQP 205
73 -AANTVELGFSP-----LGKYLSTWERPGKADGTQPKNMKVMNTERGQLVFSPVQRNQT 126
206 GWRREISVCGDVTTLAETRSAGQKLVESSTNVLDQSLDLCATLTLMTADGLFHTP 265
127 G-WNL-----QYTCDESLAA-----RLVTNVEHFEYETGNMSKGIADLRVEGISDFALSP 175
266 TOKHIEALROEINARPOCPVGLNTLAFPSIN 297
176 GQNHAVAV-----FIPKKGAPASVATTSIENFN 204

RESULT 12
AI2017
hypothetical protein all1695 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AI2017
R:Kaneho, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iritiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Ref: 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AI2017
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1449 <RUR>
A:Cross-references: UNIPROT:Q8YWC0; UNIPARC:UP100000CE194; GB:BA000019; PIDN:BA078061.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all1695

Query Match 4.0%; Score 91.5; DB 2; Length 1449;
Best Local Similarity 22.5%; Pred. No. 72;
Matches 74; Conservative 56; Mismatches 126; Indels 73; Gaps 19;

12 PNEKPYKGYELVLTGNGALPNDGRGRKSRFALYKPKANGVCPST---VHVIISTQAS 68
44 PNVAVAVY-----KNQKITYOLEKNSQNFAYTLR--HIGVKKETLVALLYERSPDVI 94
69 KAI-----SCKGQH---SISYTLRNQTVVEYTHDKTDMFQVGRSTESPIDV---VTD 119
95 IAILAILAKAGAYIPLDISAPLELEHTI-----KNAEAF-IITDESQNTYIEKTTDE 147
120 ISSGQNTDEAQTOSTISRPACRIYCDRNEPYTAIRFAPGDSKNIFLGEKAKAKNPD 179
148 IQTICIDNQNLTDFSDISLCEVTAMNLAVV---MTSG-----ST 187
180 GHMDGLTT---NGVLVMPHPRGFTESQPGVWKEI-SVCGDVTTLAETRSAGQKLVES 236
188 GKPKGVCEHGRVRLVKNNTNYANFSCEVLIQASIAFADATEIWAALLNGKLVIMP 247
237 TNN--LQD--GSLDLCATLTLMTADGLFHTPTQ---KHIEALROEINARPOCPVGLNT 290
248 INPSLQELIGALIKQYHTVITMLTA--GLFNLMVEQIHLKSLKQL-----AGGDV 298

QY 291 LAFPSINRKEVEKQPMAYISCGHVG 319
DB 299 LSVYHWSK--VIEELP-----NCQLING 320

RESULT 13
B33830
cation efflux system membrane protein czcb - Alcaligenes eutrophus

C:Species: Alcaligenes eutrophus
C:Date: 23-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004
R:Accession: B33830
R:Nies, D.H.; Nies, A.; Chu, L.; Silver, S.
Proc. Natl. Acad. Sci. U.S.A. 86, 7351-7355, 1989
A:Title: Expression and nucleotide sequence of a plasmid-determined divalent cation efflux system membrane protein; lipoyl/biotin-binding homology #status atypical <LPB>
A:Reference number: A33830; MUID:90017477; PMID:2678100
A:Accession: B33830
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-520 <NIB>
A:Cross-references: UNIPROT:P13510; UNIPARC:UP10000128D83
C:Superfamily: cation efflux system membrane protein; lipoyl/biotin-binding homology
F;217-260,333-362/Domain: lipoyl/biotin-binding homology #status atypical <LPB>

Query Match 4.0%; Score 90.5; DB 2; Length 520;
Best Local Similarity 23.3%; Pred. No. 21;
Matches 53; Conservative 23; Mismatches 94; Indels 57; Gaps 8;

QY 94 THDKDQTMFQGRSTESPIDVVTDTISGQNTDEAQTOSTISRPACRIYCDRNEP--- 150
DB 277 TYREKKLWEGKISAQ-----DYLARNALQEAQISQVNAQKLTALGASNSTALN 329
QY 151 -YTAIRFAPGDSKNIFLGEKAKAKNPDGMDGLTTNGVLVMPHPRGFTESQPGVWR 209
DB 330 RYELRAPFAGMIVEKHSIGEAADVADNANV-----FTLSDLSSVMA 369
QY 210 EISVCG-DVYTLR--ETRSAGQKLVESSTNVLDQSL-----IDLQATLW 255
DB 370 EFVSAKDVERRVIGKASINSNSDVKADGVTSYVGSLLGEQRTAKARVTLTPQMAV 429
QY 256 RTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAFPSINKEV 302
DB 430 R--PGLFVT-----VDVFGADVEVPVAVKTAVDVNGESV 464

RESULT 14
P90914
hypothetical protein Ece2286 [imported] - Escherichia coli (strain O157:H7, substrain RIN

C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: P90914
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gaswaha, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Ref: 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome
A:Reference number: A9629; MUID:21156231; PMID:11258796
A:Accession: P90914
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-823 <HAY>
A:Cross-references: UNIPROT:Q8X280; UNIPARC:UP100000D0F1E; GB:BA000007; PIDN:BA035709.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: Ece2286

Query Match 4.0%; Score 90.5; DB 2; Length 823;
Best Local Similarity 21.2%; Pred. No. 40;
Matches 77; Conservative 39; Mismatches 126; Indels 121; Gaps 20;

QY 10 CA--PNEKPYKGYELVLTGNGALPNDGRGRGRF---ALYKRPANGVCPSTHVIIST 64
DB 7 CAALPDEQALKEGAVAVA--TAIEAGDERRAAKHMQCLEHYPPAADC----- 54

QY 65 PQASKAISCCKGSHISYTLSTSRNQTVVVEYTHDKDTDFQVGRSTESPIFVVTDTISGSQ 124
 Db 55 --AYKFLVCDKXGIRPPA-----LDSDAEVWQNRNDEMSASVVPVETESDPM 102
 QY 125 NTDEAQTOSTISRPAFCRIYCDRNEPYTARIFAAGFDSKNI-----FLGEKAK 174
 Db 103 N-----VFEDMLAPEVQNAVWVKPFTCENTITVDWYISAQELQIEMMAT 145
 QY 175 WKNPDCG-----MDGLTTNV-----LVWPRGGFTBESQPGV-WREISVCGDVYTLRETRS 225
 Db 146 F---DGIYVALKMPEVNMVPELKLHAIQWVGHKKCPAKMPELQAEKRIWKRRGE 202
 QY 226 AQRGKLVSESETVWLODGSILDCGATLWRTADGLFHTPTQKHIEALRQEIINAARPOCP 285
 Db 203 RKAGKKT-----SVVDLARA-----RANQCHTESSTGKIN-----P 234
 QY 286 VGLNTLAPSPINRKEVEEKOPW-----AY-----LSCGHVHG-YHWGHRSDTEAN 331
 Db 235 V-----IAAIHR-----EYKQWTWTLDDDELALAYALMPGDVADAGNIDGSIHRMA-KNEVYDN 283
 QY 332 ERE 334
 Db 284 GRE 286

RESULT 15
 NBHDS
 Complement factor H precursor, short splice form [validated] - human
 N/Alternate names: complement factor H-related protein; complement protein H
 C/Species: Homo sapiens (hmn)
 C/Date: 31-Dec-1993 #sequence revision 23-Feb-1996 #text change 09-Jul-2004
 C/Accession: S03013; B60238; A27877; A61103; A26505; S10479
 R/Ritpoche, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.
 Biochem. J. 249, 593-602, 1988
 A/Title: The complete amino acid sequence of human complement factor H.
 A/Reference number: S00254; MUID:88134059; PMID:2963625
 A/Accession: S03013
 A/Molecule type: mRNA
 A/Residues: 1-449 <RTP>
 A/Residues: UNIPROT:P08603; UNIPARC:UPI000002635; EMBL:X07523; EMBL:Y00716; NID:A/Note: part of this sequence, including the amino end of the mature protein was conflict
 A/Note: 402-Tyr was also found
 R/Schaller, C.; Schwaebler, W.; Dietrich, M.; Weiss, E.H.
 Eur. J. Immunol. 21, 799-802, 1991
 A/Title: Human complement factor H: two factor H proteins are derived from alternatively
 A/Reference number: A60238; MUID:91184292; PMID:1826264
 A/Accession: B60238
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-33; 434-449 <EST>
 A/Cross-references: UNIPARC:UPI0000174288; UNIPARC:UPI0000174289
 A/Note: only portions of this 1.8 kilobase mRNA were sequenced
 R/Schulz, T.F.; Schwaebler, W.; Stanley, K.K.; Weiss, E.; Dietrich, M.P.
 Eur. J. Immunol. 16, 1351-1355, 1986
 A/Title: Human complement factor H: isolation of cDNA clones and partial cDNA sequence c
 A/Reference number: A27877; MUID:87054207; PMID:2946589
 A/Accession: A27877
 A/Molecule type: mRNA
 A/Residues: 1-55; 401, 'Y', 403-449 <SCH>
 A/Cross-references: UNIPARC:UPI000016A08; GB:X04697; NID:931991; PIDN:CAB41739.1; PID:R/Note: an additional nucleotide present within the codon for Glu-310 was thought to be
 R/Schwaebler, W.; Zwitter, J.; Schulz, T.F.; Linke, R.P.; Dietrich, M.P.; Weiss, E.H.
 Eur. J. Immunol. 17, 1485-1489, 1987
 A/Title: Human complement factor H: expression of an additional truncated gene product c
 A/Reference number: A61103; MUID:88055295; PMID:2445583
 A/Accession: A61103
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 27-76 <SC2>
 A/Cross-references: UNIPARC:UPI00001742EA
 A/Note: this is a partial sequence of an alternatively spliced 1.8 kilobase mRNA that is
 R/Sim, R.B.; Discolo, R.G.

Biochem. J. 205, 285-293, 1982
 A/Title: Purification and structural studies on the complement-system control protein bet
 A/Reference number: A26505; MUID:83048213; PMID:6215918
 A/Accession: A26505
 A/Molecule type: protein
 A/Residues: 19-20, 'O', 22-29, 'V', 31-33, 'Q', 35 <SIM>
 A/Cross-references: UNIPARC:UPI00001742B5
 R/Bartlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C.;
 Biochemistry 31, 3626-3634, 1992
 A/Title: Solution structure of the fifth repeat of factor H: A second example of the com
 A/Reference number: A44551; MUID:92232649; PMID:1533152
 A/Contents: annotation; NMR structure determination, residues 264-292
 R/Kristensen, T.; Wetzel, R.A.; Tack, B.F.
 J. Immunol. 136, 3407-3411, 1986
 A/Title: Structural analysis of human complement protein H: homology with C4b binding pr
 A/Reference number: S10479; MUID:86169701; PMID:2537845
 A/Accession: S10479
 A/Molecule type: mRNA
 A/Residues: 226-401, 'Y', 403-449 <KRI>
 A/Cross-references: UNIPARC:UPI000016A6C8; GB:M12383; NID:9180472; PIDN:AAA52013.1; PID:5
 C/Comment: Factor H has also been found bound to cell membranes in an unknown manner. Ho
 C/Comment: Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liver.
 C/Genetics: <HF1>
 A/Gene: GDB:HF1; HF
 A/Cross-references: GDB:120041; OMIM:134370
 A/Map position: 1q32-1q32
 C/Genetics: <HF2>
 A/Gene: GDB:HF2; HF
 A/Cross-references: GDB:129095
 A/Map position: 1q32-1q32
 A/Note: the correspondence between the two loci and the sequences indicated is unclear; f
 C/Function:
 A/Description: a cofactor in the inactivation of C3b by serine proteinase I; also increas
 he alternative complement pathway
 A/Pathway: complement alternate pathway
 C/Superfamily: complement factor H; complement factor H repeat homology
 C/Keywords: alternative splicing; complement alternate pathway; glycoprotein; plasma
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-449/Product: complement factor H, short splice form #status experimental <MAT>
 F:21-80/Domain: complement factor H repeat homology <FH01>
 F:85-141/Domain: complement factor H repeat homology <FH02>
 F:146-205/Domain: complement factor H repeat homology <FH03>
 F:210-262/Domain: complement factor H repeat homology <FH04>
 F:246-248/Region: cell attachment (R-G-D) motif
 F:267-320/Domain: complement factor H repeat homology <FH05>
 F:325-385/Domain: complement factor H repeat homology <FH06>
 F:389-442/Domain: complement factor H repeat homology <FH07>
 F:21-66, 52-80, 85-129, 114-141, 146-192, 178-205, 210-251, 237-262, 267-309, 294-320, 325-374, 357-
 F:217/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 3.9%; Score 90; DB 1; Length 449;
 Best Local Similarity 20.1%; Pred. No. 19;
 Matches 67; Conservative 37; Mismatches 124; Indels 106; Gaps 15;

QY 117 TDTISGQNTDEAQTOSTISRPAFCRIYCDRNEPYTARIFAAGFDSKNIPLGEKAKW-175
 Db 30 TELTGSWS--DQTYREGQAIYKCR-----EYRSLGNVIMCRGENV 72

QY 176 -----KNPDGNDGLTTNGVLVWHPRG-----FTEE 202
 Db 73 ALNPLRKQGRPCGH--GDPDPFGFTPL--TGGVVFYGVAVYTCNEGYYQLGEINVRG 129

QY 203 SQGVKREISVCGDVYTLAETRSQAQRGLVSEETN-----VLQDGSILDC 249
 Db 130 DTGDMTNDIPICVNVCLPVT--APENGKIVSSAMBPDRYHFGQAVRFVCSNGYKIE-- 195

QY 250 GATILWRTADGLFHTPTQKHIEALRQEIINAARPOCPVGLNTLAPSPINRKEVEEKOPMA 309
 Db 186 GDEPMKSDGDFGSKKPKCV-----EICKSPDIVNG-----SPISQKIYKEMERFQ 234

QY 310 YLSCGHVHGTHWGHRSDEANERCPMCRV--GPYPV-----LWTG 350
 Db 235 Y-KCMNGEYSEKGDVAVCTESSGWRPLPSCBEKSCDNPIYIPNGDYSPLRIKHRTGDEITVQ 293

OY 351 CEAGFYVDAGPPTTHAFTPCGHV----CSEKSAXY 380
Db 294 CRNGFIIPATRGNTAKCTSTGMIIPARCTLKPCTY 327

Search completed: March 1, 2006, 21:00:48
Job time : 42 secs

This Page Blank (uspto)

this page blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model1

Run on: March 1, 2006, 20:52:46 ; Search time 185 Seconds
(without alignments)
997.509 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MFSPGQEHGAPNKPVKYK.....ATQVGRQNCIKLIFGSPID 420

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq.21:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2290	100.0	420	5	ABB78077 Amino aci
2	2290	100.0	420	5	AAO17522 Human M33
3	2290	100.0	420	6	ABR41083 Human MAP
4	2290	100.0	420	7	ADC31180 Human nov
5	2283	99.7	420	5	ABB07922 Human pel
6	2283	99.7	420	5	ADP48671 Human pel
7	2175.5	95.0	419	5	ABB07921 Murine pe
8	2175.5	95.0	419	8	ADP48669 Mouse pel
9	2111.5	92.2	419	6	ABR41082 Mouse MAP
10	1917	83.7	418	5	ABB07920 Human pel
11	1917	83.7	418	5	AAO17518 Human M30
12	1917	83.7	418	6	ABU00017 Human nov
13	1917	83.7	418	8	ADP83103 Human pel
14	1917	83.7	418	8	ADP44005 Amino aci
15	1917	83.7	418	8	ADP48667 Human pel
16	1917	83.7	418	8	ADP56420 Human PRO
17	1917	83.7	418	8	ADP24929 PRO polyp
18	1917	83.7	418	9	ADY28614 PRO polyp
19	1917	83.7	418	9	ADY17390 PRO polyp
20	1917	83.7	418	9	AAO17515 Rat M30 p
21	1912	83.5	418	5	ABB78078 Amino aci
22	1912	83.5	418	5	ABB07919 Murine pe
23	1912	83.5	418	8	ADP48665 Mouse pel
24	1910	83.4	418	5	AAO17519 Human M30

25	1642	71.7	445	5	ABB07923 Human pel
26	1642	71.7	445	7	ADD71147 Human int
27	1642	71.7	445	8	ADP48675 Human pel
28	1640	71.6	445	5	AAO17521 Human M31
29	1633	71.3	445	7	ADC31807 Human nov
30	1632	71.3	445	5	AAO17520 Murine M3
31	1582.5	69.1	415	8	ADP44009 Amino aci
32	1492	65.2	320	5	AAO17516 Human M30
33	1492	65.2	320	5	AAO17517 Human M30
34	1482.5	64.7	406	7	ADP85774 Human pro
35	1482	64.7	320	4	AAO17522 Human pro
36	1482	64.7	320	4	AAO17523 Human pro
37	1302	56.9	424	4	AAO17524 Human pro
38	1302	56.9	424	5	AAO17525 Human pro
39	1302	56.9	424	8	ADP48676 Human pro
40	1300	56.8	320	4	AAO17526 Human pro
41	1300	56.8	320	8	ADL31994 Human pro
42	1235	53.9	455	5	ABB07925 Ascidian
43	1235	53.9	455	8	ADP48677 Clona int
44	1102	48.1	196	7	ADC32882 Human nov
45	1077	47.0	241	4	ABG04494 Novel hum

ALIGNMENTS

RESULT 1
ID ABB78077 standard; protein; 420 AA.
XX
AC ABB78077;
DT 05-NOV-2002 (first entry)
XX
XX Amino acid sequence of human Pellino 2.
DE
XX
XX Pellino 2; cancer; Pellino 1; cancer treatment; epithelial cancer;
KW gastrointestinal tract cancer.
OS
XX Homo sapiens.
XX
XX WO200259641-A2.
XX
XX 01-PAG-2002.
XX
XX 28-DEC-2001; 2001WO-US051368.
XX
XX 02-JAN-2001; 2001US-0259502P.
XX
XX (TULAR) TULARIK INC.
XX
XX Powers S, Mu D, Xiang P, Peng Y;
XX
XX WPI: 2002-619185/66.
XX
XX N-PSDB; ABQ78319.
XX
XX Detecting cancer cells in mammalian sample, useful for identifying
XX inhibitors for treating cancer e.g. epithelial cancer, comprises
XX detecting an overexpression of, or increase in copy number of genes
XX encoding, Pellino 1 and Pellino 2.
XX
XX Claim 35; Page 57-61; 69pp; English.
XX
XX The present sequence represents Pellino 2. The specification describes a
XX method for detecting cancer cells in biological sample from a mammal. The
XX method comprises detecting an overexpression of, or increase in copy
XX number of genes encoding, polypeptides Pellino 1 or Pellino 2. The method
XX is useful in detecting cancer or propensity to develop cancer, monitoring
XX the efficacy of cancer treatment, identifying inhibitors of Pellino 1 and
XX 2, inhibiting the expression and/or activity of Pellino 1 and 2 in cancer
XX cells, and treating cancer or inhibiting proliferation of cancer. The
XX cancer can be epithelial cancer, such as lung, colon, ovarian, breast,
XX prostate, kidney, stomach, bladder, or any cancer of the gastrointestinal

CC tract
XX
SQ Sequence 420 AA;

Query Match 100.0%; Score 2290; DB 5; Length 420;
Best Local Similarity 100.0%; Pred. No. 2.8e-226;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSBQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
DB 1 MFSBQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKDTMPQVGRSTSPIDFVVTDTI 120
DB 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKDTMPQVGRSTSPIDFVVTDTI 120
QY 121 SSGSNTDEAQTOSTISRACRIYCDRNEPYTARIPAGDSSKNIFLGEKAKMKNPDG 180
DB 121 SSGSNTDEAQTOSTISRACRIYCDRNEPYTARIPAGDSSKNIFLGEKAKMKNPDG 180
QY 181 HMDGLTNGVLVMPHPRGPFTEESOPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240
DB 181 HMDGLTNGVLVMPHPRGPFTEESOPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240
QY 241 QDGSLLDLCGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
DB 241 QDGSLLDLCGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
QY 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMTGCEAGFYVDAG 360
DB 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMTGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHAFHACPCATQLVGEONCIKLIPOGPID 420
DB 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHAFHACPCATQLVGEONCIKLIPOGPID 420

RESULT 2

AA017522 ID AA017522 standard; protein; 420 AA.

XX AA017522;

DT 19-JUL-2002 (first entry)

XX Human M33 protein SEQ ID NO: 16.

KW Neurodegenerative disease; M30; M31; M32; M33; stroke;
KW fragile X syndrome; Huntington's disease; Parkinson's disease;
KW Alzheimer's disease; multiple sclerosis; ovarian cancer;
KW neurodegeneration; immune disorder; autoimmune disease; allergy;
KW infection; leukemia; inflammation; neuroprotective; cerebroprotective;
KW immunosuppressive; cytotoxic; motropic; antiparkinsonian; antiallergic;
KW virucide; antiinflammatory.

XX Homo sapiens.

XX W0200221138-A2.

XX 14-MAR-2002.

XX 07-SEP-2001; 2001MO-EP010366.

XX 07-SEP-2000; 2000US-00657479.

XX (AXAR-) AXARON BIOSCIENCE AG.

XX Schneider A, Hiemisch H, Rosner M, Klugmann M, Naim J;

XX Eisenhardt G, Kuner R, Lanahan A, Worley P, Spielvogel D, Scheek S;

XX WPI, 2002-292287/33.

XX N-PSDB; AAL46323.

PT Diagnosis of neurodegenerative disease comprises detecting level of M30-
PT family proteins.

XX Claim 2; Page 116-117; 130pp; German.

CC The present invention relates to a method of diagnosing neurodegenerative
CC diseases, comprising determining the concentration of a protein in a body
CC sample, where the protein may be M30 or a variant thereof, M31, M32 or
CC M33. The method is used to diagnose neurodegenerative diseases,
CC particularly stroke but also e.g. fragile X syndrome, Huntington's,
CC Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
CC overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
CC especially ovarian cancer. The proteins can be used to identify specific
CC ligands, potentially useful for treating neurodegeneration, immune-system
CC disorders (e.g. autoimmune diseases, allergy, viral infection, leukemia,
CC inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
CC between the proteins and the protein kinase IRAK-1 can be used to treat
CC neurodegeneration. The present sequence is a protein used in the method
CC of the invention

XX Sequence 420 AA;

Query Match 100.0%; Score 2290; DB 5; Length 420;
Best Local Similarity 100.0%; Pred. No. 2.8e-226;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSBQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
DB 1 MFSBQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKDTMPQVGRSTSPIDFVVTDTI 120
DB 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKDTMPQVGRSTSPIDFVVTDTI 120
QY 121 SSGSNTDEAQTOSTISRACRIYCDRNEPYTARIPAGDSSKNIFLGEKAKMKNPDG 180
DB 121 SSGSNTDEAQTOSTISRACRIYCDRNEPYTARIPAGDSSKNIFLGEKAKMKNPDG 180
QY 181 HMDGLTNGVLVMPHPRGPFTEESOPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240
DB 181 HMDGLTNGVLVMPHPRGPFTEESOPGVWREISVCGDYTLRETRSAQORGLVSESTNVL 240
QY 241 QDGSLLDLCGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
DB 241 QDGSLLDLCGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
QY 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMTGCEAGFYVDAG 360
DB 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMTGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHAFHACPCATQLVGEONCIKLIPOGPID 420
DB 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHAFHACPCATQLVGEONCIKLIPOGPID 420

RESULT 3

ABR41083 ID ABR41083 standard; protein; 420 AA.

XX ABR41083;

XX 22-MAY-2003 (first entry)

XX Human MAP kinase cascade activator #59.

KW Human; Elki phosphorylation; Elki phosphorylation kinase; virucide;
KW antiinflammatory; immunomodulator; cytotoxic; antiallergic; anti-HIV;
KW antineumatic; antiarthritic; antidiabetic; antiaesthetic; gene therapy;
KW inflammation; autoimmune disease; viral disease; cancer; diabetes;
KW rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
KW Iga nephritis.

XX Homo sapiens.

XX WO2003008589-A1.
 XX 30-JAN-2003.
 XX 15-JUL-2002; 2002WO-JP007174.
 XX 18-JUL-2001; 2001JP-00218204.
 XX 31-AUG-2001; 2001JP-00263450.
 XX 21-JAN-2002; 2002JP-00012176.
 XX (ASAH) ASAH KASEI KOGYO KK.
 XX Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
 XX WPI, 2003-229582/22.
 XX N-PSDB; ACC42349.
 XX Etk1 phosphorylation-associated gene and its encoded protein with MAP
 XX kinase cascade effect, applicable in diagnosis of and developing drugs
 XX for e.g. inflammation, autoimmune diseases, viral diseases and cancer.
 XX Claim 1; Page 640-643; 762pp; Japanese.
 XX The invention relates to a novel purified protein having Etk1
 XX phosphorylation activity and/or an activity of activating Etk1
 XX phosphotransferase kinase. A protein of the invention has anti-inflammatory,
 XX immunomodulator, virucide, cytostatic, antiallergic, antineoplastic,
 XX antitumor, antidiabetic, antiasthmatic, and anti-HIV activity. The
 XX polynucleotides may have a use in gene therapy. The gene and its encoded
 XX protein are applicable in diagnosis of and developing drugs for e.g. '
 XX inflammations, autoimmune diseases, viral diseases and cancer such as
 XX rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral
 XX hepatitis and Iga nephritis. The present sequence is used in the
 XX exemplification of the invention.
 XX Sequence 420 AA:
 SQ
 Query Match 100.0%; Score 2290; DB 6; Length 420;
 Best Local Similarity 100.0%; Pred. No. 2.8e-226;
 Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MFSRQGEHCAKNEPKYKGLVYVLYGNGALPNDGRGRKSRFLYRPRKANGVKESTVH 60
 DB 1 MFSRQGEHCAKNEPKYKGLVYVLYGNGALPNDGRGRKSRFLYRPRKANGVKESTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDQTMFQVGRSTESPIDVVTDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDQTMFQVGRSTESPIDVVTDTI 120
 QY 121 SSGQNTDEAOITOSTISRACRIVCNRPYTAIRIFAGPDSKNIFLGEKAAKKNPDG 180
 DB 121 SSGQNTDEAOITOSTISRACRIVCNRPYTAIRIFAGPDSKNIFLGEKAAKKNPDG 180
 QY 181 HMOGLTNGVLYVMPHPRGFTFESQPVWREISVCGDVYTLRETSSAQRGLVSESTNVL 240
 DB 181 HMOGLTNGVLYVMPHPRGFTFESQPVWREISVCGDVYTLRETSSAQRGLVSESTNVL 240
 QY 241 QDSGLIDLCATLTMTADGLFHTPTQKHTALQOEINAAARPOCPVGLNTLAFSINRKE 300
 DB 241 QDSGLIDLCATLTMTADGLFHTPTQKHTALQOEINAAARPOCPVGLNTLAFSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGHYNWGRSDTEANERECPMCRITGVPPVPLMLGCEAGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGHYNWGRSDTEANERECPMCRITGVPPVPLMLGCEAGFYVDAG 360
 QY 361 PPTHTAFTPCGHVCEKSAKXWSQIPLPHGTHAFAACPCATOLVGBONCIKILFQSPID 420
 DB 361 PPTHTAFTPCGHVCEKSAKXWSQIPLPHGTHAFAACPCATOLVGBONCIKILFQSPID 420

ID ADC31180 standard; protein; 420 AA.
 XX
 AC ADC31180;
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel polypeptide sequence, SEQ ID NO:1262.
 XX
 KW Human, diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; anti-nausea; anticoagulant; thrombolytic; vulnery;
 KW antidiabetic; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 14q21.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029271-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 24-SEP-2002; 2002WO-US030474.
 XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 PA (HYSEQ-) HYSEQ INC.
 XX
 PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX
 DR WPI, 2003-371981/35.
 DR N-PSDB; ADC30209.
 XX
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 PS Claim 20; SEQ ID NO 1262; 1185pp; English.
 XX
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31850). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC cDNA sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the cDNAs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but

CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 420 AA;

Query Match 100.0%; Score 2290; DB 7; Length 420;
 Best Local Similarity 100.0%; Pred. No. 2.8e-226;
 Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESPGEHCAPKEPVKYGELVLTGYNGALPNDRGRKRSFPALYKRPKANGVKPSTVH 60
 DB 1 MESPGEHCAPKEPVKYGELVLTGYNGALPNDRGRKRSFPALYKRPKANGVKPSTVH 60
 QY 61 VISTPQASKAISCKGSHSYTLSSRNQTVVEYTHDKDMDQVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGSHSYTLSSRNQTVVEYTHDKDMDQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQITOSTISRFACRIVCDRNEPYTARIIPAAGFDSKNIFLGEKAKMKNPDG 180
 DB 121 SSGQNTDEAQITOSTISRFACRIVCDRNEPYTARIIPAAGFDSKNIFLGEKAKMKNPDG 180
 QY 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDYTLRETRSAQQRKLVESSTNVL 240
 DB 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDYTLRETRSAQQRKLVESSTNVL 240
 QY 241 QDGSLLDLCGATLLMRTADGLFHTPTQKHIEALROBINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDGSLLDLCGATLLMRTADGLFHTPTQKHIEALROBINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGHYNNWGHRSDEANERECPCRTVGYPVPLMLGCEAGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGHYNNWGHRSDEANERECPCRTVGYPVPLMLGCEAGFYVDAG 360
 QY 361 PPTHAFPCGHVCSKSAKYWSQIPLPHGTTHAFHAACPFCATQVGEQNCIKILIFQGPID 420
 DB 361 PPTHAFPCGHVCSKSAKYWSQIPLPHGTTHAFHAACPFCATQVGEQNCIKILIFQGPID 420

RESULT 5

ABB07922 standard; protein; 420 AA.

AC ABB07922;
 DT 30-JUL-2002 (first entry)
 DE Human pellino-2 polypeptide.

XX Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; vitruclide;
 KW antibacterial; fungicide; protozoan; antidiabetic; antirheumatic;
 KM antiarthritic; antiinflammatory; antiatherosclerotic; neuroprotective;
 KM nootropic; antidiabetic; human; pellino-2.

OS Homo sapiens.

XX MO200183739-A2.

PN 08-NOV-2001.

XX 27-APR-2001; 2001WO-US013676.

XX 28-APR-2000; 2000US-0200198P.

XX (IMMV) IMMUNEX CORP.

XX Bird TA, Cosman DJ;

XX WPI; 2002-066532/09.

XX N-PDB; ABL58449.

PT New Pellino polypeptides for identifying compounds that alter polypeptide
 PT activity, treating pathogenic infection or inhibiting apoptosis, are
 PT capable of stimulating nuclear factor-kappaB- or p38-dependent

PT transcription.

XX Claim 1; Page 61-63; 70pp; English.

CC The invention provides polypeptides capable of stimulating nuclear factor
 CC (NF)-kappaB-dependent transcription or p38-dependent transcription,
 CC referred as Pellino polypeptides. The pellino polypeptides are useful for
 CC identifying modulators that alter the pellino polypeptide and pellino
 CC dominant-negative activity. They are also useful for identifying
 CC compounds that inhibit the binding activity of the polypeptides and to
 CC study cell-signal transduction. They are useful for preventing or
 CC treating infection by a pathogen such as virus, bacterial, fungi, algae
 CC or protozoa, or inhibiting apoptosis. Dominant-negative pellino
 CC polypeptides are useful for treating inflammatory conditions such as
 CC asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and
 CC also for inhibiting mitogen activated protein (MAP) kinase-activated
 CC pathways. Pellino polypeptides and polynucleotides are useful to identify
 CC small molecule inhibitors of protein association or function of Pellino,
 CC and other molecules involved in interleukin (IL)-1 signaling. The present
 CC sequence represents a human pellino-2 polypeptide

XX Sequence 420 AA;

Query Match 99.7%; Score 2283; DB 5; Length 420;
 Best Local Similarity 99.8%; Pred. No. 1.5e-225;
 Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESPGEHCAPKEPVKYGELVLTGYNGALPNDRGRKRSFPALYKRPKANGVKPSTVH 60
 DB 1 MESPGEHCAPKEPVKYGELVLTGYNGALPNDRGRKRSFPALYKRPKANGVKPSTVH 60
 QY 61 VISTPQASKAISCKGSHSYTLSSRNQTVVEYTHDKDMDQVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGSHSYTLSSRNQTVVEYTHDKDMDQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQITOSTISRFACRIVCDRNEPYTARIIPAAGFDSKNIFLGEKAKMKNPDG 180
 DB 121 SSGQNTDEAQITOSTISRFACRIVCDRNEPYTARIIPAAGFDSKNIFLGEKAKMKNPDG 180
 QY 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDYTLRETRSAQQRKLVESSTNVL 240
 DB 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDYTLRETRSAQQRKLVESSTNVL 240
 QY 241 QDGSLLDLCGATLLMRTADGLFHTPTQKHIEALROBINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDGSLLDLCGATLLMRTADGLFHTPTQKHIEALROBINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGHYNNWGHRSDEANERECPCRTVGYPVPLMLGCEAGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGHYNNWGHRSDEANERECPCRTVGYPVPLMLGCEAGFYVDAG 360
 QY 361 PPTHAFPCGHVCSKSAKYWSQIPLPHGTTHAFHAACPFCATQVGEQNCIKILIFQGPID 420
 DB 361 PPTHAFPCGHVCSKSAKYWSQIPLPHGTTHAFHAACPFCATQVGEQNCIKILIFQGPID 420

RESULT 6

ADP48671 standard; protein; 420 AA.

AC ADP48671;

XX 09-SEP-2004 (first entry)

XX Human Pellino-2 protein SEQ ID NO:8.

XX nuclear factor kappa B dependent transcription inhibitor;
 KW NF-kB-dependent transcription inhibitor;
 KW p38-dependent transcription inhibitor; Pellino-1;
 KW interleukin 1 receptor-associated kinase 4;
 KW IL-1 receptor-associated kinase 4; IRAK-4; antiatherosclerotic;
 KW antiarthritic; antidiabetic; antiinflammatory; antirheumatic; antidiabetic;

KW gastrointestinal; neuroprotective; nootropic; IL-1 antagonist;
 KW IRAK protein kinase family inhibitor;
 KW NF kappa B inducing kinase inhibitor; asthma; rheumatoid arthritis;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW atherosclerosis; Alzheimer's disease; human; Pellino-2; chromosome 14.
 XX
 OS Homo sapiens.
 XX
 WO2004053092-A2.
 XX
 PN 24-JUN-2004.
 XX
 PD 09-DEC-2003; 2003WO-US039188.
 XX
 PF 11-DEC-2002; 2002US-00317250.
 XX
 PR (IMMV) IMMUNEX CORP.
 XX (CLEV) CLEVELAND CLINIC FOUND.
 PA
 XX
 PI Bird TA, Cosman DJ, Li X;
 XX
 XX WPI; 2004-480927/45.
 DR N-PSDB; ADP48670.
 XX
 PT Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
 PT asthma, and rheumatoid arthritis, comprises assaying the association of a
 PT Pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
 XX
 PS Example 1; SEQ ID NO 8; 81pp; English.

50 The present invention describes a method for identifying compounds (C) that inhibit nuclear factor kappa B (NF- κ B)-dependent transcription or p38-dependent transcription. The method comprises: (1) mixing a test compound with a Pellino-1 polypeptide; (b) assaying the association of the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4) in the presence of the test compound; and (c) determining whether the test compound inhibits the association of the Pellino-1 polypeptide with a binding partner. Also described: (1) an inhibitory nucleic acid that binds to a nucleic acid encoding an amino acid sequence as described above, where the presence of the inhibitory nucleic acid within a cell inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory polypeptide comprising an antibody fragment that binds to a polypeptide comprising an amino acid sequence as described above, where the presence of the inhibitory polypeptide within a cell inhibits the association of Pellino-1 with IRAK-4. (C) have antiatherosclerotic, antiarthritic, antiasthmatic, antiinflammatory, antineurotic, antitumor, gastrointestional, neuroprotective and nociceptive activities, and can be used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF kappa B inducing kinase inhibitor. The methods and compositions of the present invention are useful for the prevention and/or treatment of diseases or conditions associated with aberrant expression or activity of the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, atherosclerosis and Alzheimer's disease. The present sequence represents human Pellino-2, which is used in the exemplification of the present invention. The human Pellino-2 gene is located on chromosome 14, more specifically to 14q24.3.

	Query Match	99.7%	Score 2283	DB 8	Length 420
	Best Local Similarity	99.8%	Pred. No. 1.5e-225		
	Matches 419	Conservative 0	Med. Matches 1	Indels 0	Gaps 0
Qy	1	MPSPGQEEHCAPNKEPVYKGGELVWLVGYNCAALPNGDGRGRKRSFALYKRPKANGVPSVTH			60
Db	1	MPSPGQEEHCAPNKEPVYKGGELVWLVGYNCAALPNGDGRGRKRSFALYKRPKANGVPSVTH			60
Qy	61	VISTPQAKSAISCCKGSHISYTLSSHNOQVWEYTHDKDTDMQVGRSTESPIDFVVTDTI			120
Db	61	VISTPQAKSAISCCKGSHISYTLSSHNOQVWEYTHDKDTDMQVGRSTESPIDFVVTDTI			120
Qy	121	SGSONTDAQITQSTISRPRACIVCDRNEPYTARIIPAAGFDSKNIIFGEKAAKWNPDG			180

Db	121	SGSQNTDAQQLTGSTISRFACIVCDRMBPYTARIFAAGFSSSKNIFLGVAKAKMNPDG	180
Qy	181	HMDGILTTGCVLMNHRGGFTESQSGWREISVCGDYTLRTTRSAOQRGKIVESTNTLV	240
Db	181	HMDGILTTGCVLMNHRGGFTESQSGWREISVCGDYTLRTTRSAOQRGKIVESTNTLV	240
Qy	241	QDGSILIDLCGATLLMRTADGLFHTPTQKHIEALROEINNAARPOCPVGLNTTLA	300
Db	241	QDGSILIDLCGATLLMRTADGLFHTPTQKHIEALROEINNAARPOCPVGLNTTLA	300
Qy	301	VVEEKQPMAYISCGVHGTHNMGHSDDPEANRRECPMCTVPPYVPLMTGCEAGFYVDAG	360
Db	301	VVEEKQPMAYISCGVHGTHNMGHSDDPEANRRECPMCTVPPYVPLMTGCEAGFYVDAG	360
Qy	361	PPTAAPFTCCGHVCSKSAKATWSQILPHGHTAFHAACPACATOLVGEONCIKILIPQPID	420
Db	361	PPTAAPFTCCGHVCSKSAKATWSQILPHGHTAFHAACPACATOLVGEONCIKILIPQPID	420

RESULT 7	
ABB07921	
ID	ABB07921 standard; protein; 419 AA.
xx	
xx	ABB07921;
AC	
xx	
DT	30-JUL-2002 (first entry)
xx	
DE	Murine pellino-2 polypeptide.

XX Nucleofactor-kappab; NF-kappab; p38; pellino; interleukin-1; virucide;
 KW antibacterial; fungicide; protozoicide; antiaslarmic; antirheumatic;
 KW antiaarthritis; antiinflammatory; antiatherosclerotic; neuroprotective;
 XX neurotopic; antitumor; mouse; pellino-2.
 XX Mus musculus.
 OS
 PM WO200183739-A2.
 PD 08-NOV-2001.
 XX
 XX 27-APR-2001; 2001WO-US013676.
 PF 28-APR-2000; 2000US-0200198R.
 PR
 PA (IMMV) IMMUNEX CORP.
 XX
 PI Bird TA, Cosman DJ;
 XX
 DR WPI, 2002-066532/09.
 DR N-P8DB; ABL58448.
 XX
 PT New pellino polypeptides for identifying compounds that alter polypeptide
 PT activity, treating pathogenic infection or inhibiting apoptosis--are
 PT capable of stimulating nuclear factor-kappab- or p38-dependent
 PT transcripion.
 XX
 XX Example; Page 59-60; 70pp; English.
 PS
 XX
 CC The invention provides polypeptides capable of stimulating nuclear factor
 CC (NF)-kappab-dependent transcripion or p38-dependent transcripion,
 CC referred as pellino polypeptides. The pellino polypeptides are useful for
 CC identifying modulators that alter the pellino polypeptide and pellino
 CC dominant-negative activity. They are also useful for identifying
 CC compounds that inhibit the binding activity of the polypeptides and to
 CC study cell-signal transduction. They are useful for preventing or
 CC treating infection by a pathogen such as virus, bacterial, fungi, algae
 CC or protozoa, or inhibiting apoptosis. Dominant-negative pellino
 CC polypeptides are useful for treating inflammatory conditions such as
 CC asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and
 CC also for inhibiting mitogen activated protein (MAP) kinase-activated
 CC pathways. Pellino polypeptides and polynucleotides are useful to identify

CC small molecule inhibitors of protein association or function of Pellino,
 CC and other molecules involved in interleukin (IL)-1 signaling. The present
 CC sequence represents a murine pellino-2 polypeptide

XX Sequence 419 AA;

Query Match 95.0%; Score 2175.5; DB 5; Length 419;
 Best Local Similarity 95.0%; Pred. No. 1,7e-214;
 Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MESPQGEHPCAPKPEVKYGEIVLVLYGNGALPNDGRKRRKSPALYKRKANGKVESTYH 60
 DB 1 MESPQGEHPCAPKPEVKYGEIVLVLYGNGALPNDGRKRRKSPALYKRKANGKVESTYH 60
 QY 61 VISTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTDMFQVGRSTESPIDFVVTDTI 120
 DB 61 MVTTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTDMFQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKKNPDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKKNPDG 179
 QY 181 HMGGLTNGVLVNHPRGFTBESQPVWREISVCGDYTTRETRSAQQRKLVESSTNVL 240
 DB 180 HMGGLTNGVLVNHPRGFTBESQPVWREISVCGDYTTRETRSAQQRKLVESSTNVL 239
 QY 241 QDGSLLDLCGATLLMTADGLFTPTOKHLEALROEINARPOCPVGLNTLAPPSINRKE 300
 DB 240 QDGSLLDLCGATLLMTADGLFTPTOKHLEALROEINARPOCPVGLNTLAPPSINRKE 299
 QY 301 VVEKQPMAYLSCGHVGYNMGHRSDEANRECEPCRTVGYPVPLMLGCEAGFYVDAG 360
 DB 300 VVEKQPMAYLSCGHVGYNMGHRSDEANRECEPCRTVGYPVPLMLGCEAGFYVDAG 359
 QY 361 PPTHAFPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCATQLVGEONCIKILFQGPID 420
 DB 360 PPTHAFPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCATQLVGEONCIKILFQGPID 419

RESULT 8

ID ADP48669 standard; protein; 419 AA.

AC ADP48669;

DT 09-SEP-2004 (first entry)

DE Mouse Pellino-2 protein SEQ ID NO:6.

XX nuclear factor kappa B dependent transcription inhibitor;
 XX NF-kB-dependent transcription inhibitor;
 XX p38-dependent transcription inhibitor; Pellino-1;
 XX interleukin 1 receptor-associated kinase 4;
 XX IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
 XX antiarthritic; antiasthmatic; antiinflammatory; antirheumatic;
 XX gastrointestinal; neuroprotective; nocotropic; IL-1 antagonist;
 XX IRAK protein kinase family inhibitor;
 XX NF kappa B inducing kinase inhibitor; asthma; rheumatoid arthritis;
 XX inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 XX atherosclerosis; Alzheimer's disease; mouse; Pellino-2.

OS Mus musculus.

PN W02004053092-A2.

XX 24-JUN-2004.

PF 09-DEC-2003; 2003MO-US039188.

PR 11-DEC-2002; 2002US-00317250.

XX (IMMUNEX CORP.

PA (CLEVELAND CLINIC FOUND.

XX Bird TA, Cosman DJ, Li X;
 PI WPI; 2004-480927/45.
 DR N-PSDB; ADP48668.

XX Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
 PT asthma, and rheumatoid arthritis, comprises assaying the association of a
 PT Pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
 XX Example 1; SEQ ID NO 6; 81pp; English.

PS The present invention describes a method for identifying compounds (C)
 XX that inhibit nuclear factor kappa B (NF-kB)-dependent transcription or
 CC p38-dependent transcription. The method comprises: (1) mixing a test
 CC compound with a Pellino-1 polypeptide; (b) assaying the association of
 CC the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4)
 CC in the presence of the test compound; and (c) determining whether the
 CC test compound inhibits the association of the Pellino-1 polypeptide with
 CC a binding partner. Also described: (1) an inhibitory nucleic acid that
 CC binds to a nucleic acid encoding an amino acid sequence as described
 CC above, where the presence of the inhibitory nucleic acid within a cell
 CC inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory
 CC polypeptide comprising an antibody fragment that binds to a polypeptide
 CC comprising an amino acid sequence as described above, where the presence
 CC of the inhibitory polypeptide within a cell inhibits the association of
 CC Pellino-1 with IRAK-4. (C) have antiarteriosclerotic, antiarthritic,
 CC antiasthmatic, antiinflammatory, antirheumatic, antitumor,
 CC gastrointestinal, neuroprotective and nocotropic activities, and can be
 CC used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF
 CC kappa B inducing kinase inhibitor. The methods and compositions of the
 CC present invention are useful for the prevention and/or treatment of
 CC diseases or conditions associated with aberrant expression or activity of
 CC the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC atherosclerosis and Alzheimer's disease. The present sequence represents
 CC mouse Pellino-2, which is used in the exemplification of the present
 CC invention.

XX Sequence 419 AA;

Query Match 95.0%; Score 2175.5; DB 8; Length 419;
 Best Local Similarity 95.0%; Pred. No. 1,7e-214;
 Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MESPQGEHPCAPKPEVKYGEIVLVLYGNGALPNDGRKRRKSPALYKRKANGKVESTYH 60
 DB 1 MESPQGEHPCAPKPEVKYGEIVLVLYGNGALPNDGRKRRKSPALYKRKANGKVESTYH 60
 QY 61 VISTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTDMFQVGRSTESPIDFVVTDTI 120
 DB 61 MVTTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTDMFQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKKNPDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKKNPDG 179
 QY 181 HMGGLTNGVLVNHPRGFTBESQPVWREISVCGDYTTRETRSAQQRKLVESSTNVL 240
 DB 180 HMGGLTNGVLVNHPRGFTBESQPVWREISVCGDYTTRETRSAQQRKLVESSTNVL 239
 QY 241 QDGSLLDLCGATLLMTADGLFTPTOKHLEALROEINARPOCPVGLNTLAPPSINRKE 300
 DB 240 QDGSLLDLCGATLLMTADGLFTPTOKHLEALROEINARPOCPVGLNTLAPPSINRKE 299
 QY 301 VVEKQPMAYLSCGHVGYNMGHRSDEANRECEPCRTVGYPVPLMLGCEAGFYVDAG 360
 DB 300 VVEKQPMAYLSCGHVGYNMGHRSDEANRECEPCRTVGYPVPLMLGCEAGFYVDAG 359
 QY 361 PPTHAFPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCATQLVGEONCIKILFQGPID 420
 DB 360 PPTHAFPCGHVSEKSAKYSQIPLPHGTHAFHAACPFCATQLVGEONCIKILFQGPID 419

RESULT 9
 ABR41082
 ID ABR41082 standard; protein; 419 AA.
 XX
 AC ABR41082;
 XX
 DT 22-MAY-2003 (first entry)
 XX
 DE Mouse MAP kinase cascade activator #9.
 XX
 KW Mouse; Etk1 phosphorylation; Etk1 phosphorylation kinase; virucide;
 KW antiinflammatory; immunomodulator; cytostatic; antiallergic; anti-HIV;
 KW antineumatic; antiarthritic; antidiabetic; antiasumatic; gene therapy;
 KW inflammation; autoimmune disease; viral disease; cancer; diabetes;
 KW rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
 KW IGA nephritis.
 XX
 OS Mus musculus.
 XX
 PN WO2003008589-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 15-JUL-2002; 2002WO-JP007174.
 XX
 PR 18-JUL-2001; 2001JP-00218204.
 PR 31-AUG-2001; 2001JP-00263450.
 PR 21-JAN-2002; 2002JP-00012176.
 XX
 PA (ASAH) ASAH KASEI KOGYO KK.
 XX
 PI Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
 XX
 DR WPI; 2003-229582/22.
 DR N-PSDB; ACC42348.
 XX
 PT Etk1 phosphorylation-associated gene and its encoded protein with MAP
 PT kinase cascade effect, applicable in diagnosis of and developing drugs
 PT for e.g. inflammation, autoimmune diseases, viral diseases and cancer.
 XX
 PS Claim 1; Page 628-630; 762pp; Japanese.
 XX
 CC The invention relates to a novel purified protein having Etk1
 CC phosphorylation activity and/or an activity of activating Etk1
 CC phosphorylation kinase. A protein of the invention has antiinflammatory,
 CC immunomodulator, virucide, cytostatic, antiallergic, antineumatic,
 CC antiarthritic, antidiabetic, antiasumatic, and anti-HIV activity. The
 CC polynucleotides may have a use in gene therapy. The gene and its encoded
 CC protein are applicable in diagnosis of and developing drugs for e.g.
 CC inflammation, autoimmune diseases, viral diseases and cancer such as
 CC rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral
 CC hepatitis and IGA nephritis. The present sequence is used in the
 CC exemplification of the invention
 XX
 SQ Sequence 419 AA.

Query Match 92.4%; Score 2111.5; DB 6; Length 419;
 Best Local Similarity 92.4%; Pred. No. 6.8e-208;
 Matches 388; Conservative 12; Mismatches 19; Indels 1; Gaps 1;

QY 1 MESPQGEHCAPEKPEYKELVYLVGNGALPNDGRKRRKSPALYRPRANGVPESTVH 60
 DB 1 MESPQGEHCAPEKPEYKELVYLVGNGALPNDGRKRRKSPALYRPRANGVPESTVH 60
 QY 61 VISTPQSKAISRGHHSITLRSQTVVEVTHDQTFQVGRSTESPIDVVDY 120
 DB 61 VISTPQSKAISRGHHSITLRSQTVVEVTHDQTFQVGRSTESPIDVVDY 120
 QY 121 SGGQNTDEAQITOSTISRFACRIVCDNEPTTARIFAPAGFDSSKNITLGEKAAKKNPDG 180
 DB 121 SGGQNTDEAQITOSTISRFACRIVCDNEPTTARIFAPAGFDSSKNITLGEKAAKKNPDG 180
 QY 121 SGGQNTDEAQITOSTISRFACRIVCDNEPTTARIFAPAGFDSSKNITLGEKAAKKNPDG 179
 DB 121 SGGQNTDEAQITOSTISRFACRIVCDNEPTTARIFAPAGFDSSKNITLGEKAAKKNPDG 179

QY 181 HMDGLTTNGVLVNHPRGFTSESPGVMREISVCGDYVTLRETRSAQORGLVSESTNVL 240
 DB 180 HMDGLTTNGVLVNHPRGFTSESPGVMREISVCGDYVTLRETRSAQORGLVSESTNVL 239
 QY 241 QDGLSLIDLGCATLIMRTADGLFPTTQKHIEALRQEINARPOCPVGLNTLAPPSINRKE 300
 DB 240 QDGLSLIDLGCATLIMRTADGLFPAHQKHIEALRQEINARPOCPVGLNTLAPPSINRKE 299
 QY 301 VVEKQPMAYLSCGHVGYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVXX 360
 DB 300 VVEKQPMAYLSCGHVGYNNMGRSDTEANERECPCMTVGPPVPLMLGCEAGFYVXX 359
 QY 361 PPTHAFFPCGHVCEKSAKYSQIPLPHGTHAFAACPFATQLVGEQNCIKLIFQGPID 420
 DB 360 XXXHVFPCGHVCEKSAKYSQIPLPHGTHAFAACPFATQLVGEQNCIKLIFQGPID 419

RESULT 10
 ABB07920
 ID ABB07920 standard; protein; 418 AA.
 XX
 AC ABB07920;
 XX
 DT 30-JUL-2002 (first entry)
 XX
 DE Human pellino-1 polypeptide.
 XX
 KW Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide;
 KW antibacterial; fungicide; protozoacide; antiasumatic; antirheumatic;
 KW antiarthritic; antineumatic; antiallergic; antidiabetic; antineumatic;
 KW neurotropic; antidiabetic; human; pellino-1.
 XX
 OS Homo sapiens.
 XX
 PN WO200183739-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US013676.
 XX
 PR 28-APR-2000; 2000US-0200198P.
 XX
 PA (IMMUNEX CORP.
 XX
 PI Bird TA, Cosman DJ;
 XX
 DR WPI; 2002-066532/09.
 DR N-PSDB; ABL58447.
 XX
 PT New pellino polypeptides for identifying compounds that alter polypeptide
 PT activity, treating pathogenic infection or inhibiting apoptosis, are
 PT capable of stimulating nuclear factor-kappaB- or p38-dependent
 PT transcription.
 XX
 PS Claim 1; Page 57-58; 70pp; English.
 XX
 CC The invention provides polypeptides capable of stimulating nuclear factor
 CC (NF)-kappaB-dependent transcription or p38-dependent transcription,
 CC referred as Pellino polypeptides. The pellino polypeptides are useful for
 CC identifying modulators that alter the pellino polypeptide and pellino
 CC dominant-negative activity. They are also useful for identifying
 CC compounds that inhibit the binding activity of the polypeptides and to
 CC study cell-signal transduction. They are useful for preventing or
 CC treating infection by a pathogen such as virus, bacterial, fungi, algae
 CC or protozoa, or inhibiting apoptosis. Dominant-negative pellino
 CC polypeptides are useful for treating inflammatory conditions such as
 CC asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and
 CC also for inhibiting mitogen activated protein (MAP) kinase-activated
 CC pathways. Pellino polypeptides and polynucleotides are useful to identify
 CC small molecule inhibitors of protein association or function of pellino.
 CC and other molecules involved in interleukin (IL)-1 signaling. The present
 CC sequence represents a human pellino-1 polypeptide

XX Sequence 418 AA; 83.7%; Score 1917; DB 5; Length 418;
 Query Match Best Local Similarity 81.4%; Pred. No. 7,4e-186; Indels 2; Gaps 1;
 Matches 342; Conservative 38; Mismatches 38; Deletions 2; Gaps 1;
 QY 1 MFSPOGEHCAPKEPVKXGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
 DB 1 MFSPOGEHCAPKEPVKXGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 58
 QY 61 VISTPOAKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFQVGRSTSPIDFVTTDTI 120
 DB 59 IACTPOAKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFQVGRSTSPIDFVTTDTI 118
 QY 121 SGSONTDEAQTOSTISRFACRIVCDBNEPYTARIAPAGDSSKNFLGKAKAKMKNPDG 180
 DB 119 PGSQNSDTSQVOSTISRFACRIVCDBNEPYTARIAPAGDSSKNFLGKAKAKMKNPDG 178
 QY 181 HMDGLTTNGVLVMPHPRGFTESOPGVWREISVCGDVYTLRETRSAOQRKLVESETNVL 240
 DB 179 QMDGLTTNGVLVMPHPRGFTESOPGVWREISVCGDVYTLRETRSAOQRKLVESETNVL 238
 QY 241 QDGSILDLGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
 DB 239 QDGSILDLGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAFPSINRKE 298
 QY 301 VVEKOPMAYLSCGHVGHYNNMGHRSDETEANERECPCRTGVPYVPLMTGCEAGFYVDAG 360
 DB 299 VVEKOPMAYLSCGHVGHYNNMGHRSDETEANERECPCRTGVPYVPLMTGCEAGFYVDAG 358
 QY 361 PPTHAFTPCGHVCSEKSAKYMSQIPLPHGTHAFHACPCFATQLVGEONCIRKLIPOGPID 420
 DB 359 PPTHAFTPCGHVCSEKSAKYMSQIPLPHGTHAFHACPCFATQLVGEONCIRKLIPOGPID 418
 RESULT 11
 AA017518 standard; protein; 418 AA.
 XX AA017518;
 DT 19-JUL-2002 (first entry)
 DE Human M30 variant C protein SEQ ID NO: 8.
 XX Neurodegenerative disease; M30; M31; M32; M33; stroke;
 KW Fragile X syndrome; Huntington's disease; Parkinson's disease;
 KW Alzheimer's disease; multiple sclerosis; ovarian cancer;
 KW neurodegeneration; immune disorder; autoimmune disease; allergy;
 KW infection; leukaemia; inflammation; neuroprotective; cerebroprotective;
 KW immunosuppressive; cytostatic; noctropic; antiparkinsonian; antiallergic;
 KW virucide; antiinflammatory.
 XX Homo sapiens.
 OS Homo sapiens.
 PN MO200221138-A2.
 PD 14-MAR-2002.
 PF 07-SEP-2001; 2001WC-EP010366.
 PR 07-SEP-2000; 2000US-00657479.
 PA (AXAR-) AXARON BIOSCIENCE AG.
 PI Schneider A, Hlemisch H, Rosner M, Klugmann M, Naim J,
 PI Eisenhardt G, Kumer R, Lanahan A, Worley P, Spielvogel D, Scheek S;
 DR MPI: 2002-292287/33.
 DR N-PSDB; AAL46319.
 PT Diagnosis of neurodegenerative disease comprises detecting level of M30-

PT family proteins.
 XX Claim 2; Page 95-97; 130pp; German.
 PS The present invention relates to a method of diagnosing neurodegenerative
 CC diseases, comprising determining the concentration of a protein in a body
 CC sample, where the protein may be M30 or a variant thereof, M31, M32 or
 CC M33. The method is used to diagnose neurodegenerative diseases,
 CC particularly stroke but also e.g. fragile X syndrome, Huntington's,
 CC Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
 CC overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
 CC especially ovarian cancer. The proteins can be used to identify specific
 CC ligands, potentially useful for treating neurodegeneration, immune-system
 CC disorders (e.g. autoimmune diseases, allergy, viral infection, leukaemia,
 CC inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
 CC between the proteins and the protein kinase IRAK-1 can be used to treat
 CC neurodegeneration. The present sequence is a protein used in the method
 CC of the invention
 CC
 SQ Sequence 418 AA;
 Query Match Best Local Similarity 83.7%; Score 1917; DB 5; Length 418;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
 QY 1 MFSPOGEHCAPKEPVKXGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
 DB 1 MFSPOGEHCAPKEPVKXGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 58
 QY 61 VISTPOAKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFQVGRSTSPIDFVTTDTI 120
 DB 59 IACTPOAKAISCKGQHSISYTLSSRNQTVVVEYTHDKDTMFQVGRSTSPIDFVTTDTI 118
 QY 121 SGSONTDEAQTOSTISRFACRIVCDBNEPYTARIAPAGDSSKNFLGKAKAKMKNPDG 180
 DB 119 PGSQNSDTSQVOSTISRFACRIVCDBNEPYTARIAPAGDSSKNFLGKAKAKMKNPDG 178
 QY 181 HMDGLTTNGVLVMPHPRGFTESOPGVWREISVCGDVYTLRETRSAOQRKLVESETNVL 240
 DB 179 QMDGLTTNGVLVMPHPRGFTESOPGVWREISVCGDVYTLRETRSAOQRKLVESETNVL 238
 QY 241 QDGSILDLGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
 DB 239 QDGSILDLGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAFPSINRKE 298
 QY 301 VVEKOPMAYLSCGHVGHYNNMGHRSDETEANERECPCRTGVPYVPLMTGCEAGFYVDAG 360
 DB 299 VVEKOPMAYLSCGHVGHYNNMGHRSDETEANERECPCRTGVPYVPLMTGCEAGFYVDAG 358
 QY 361 PPTHAFTPCGHVCSEKSAKYMSQIPLPHGTHAFHACPCFATQLVGEONCIRKLIPOGPID 420
 DB 359 PPTHAFTPCGHVCSEKSAKYMSQIPLPHGTHAFHACPCFATQLVGEONCIRKLIPOGPID 418
 RESULT 12
 ABU00017 standard; protein; 418 AA.
 ID ABU00017
 AC ABU00017;
 DT 17-JAN-2003 (first entry)
 DE Human novel polypeptide #110.
 XX Human; genetic disorder; gene mapping; medical imaging; cancer;
 KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
 KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
 KW osteoarthritis; peridontal disease; liver fibrosis; viral infection;
 KW fungal infection; bacterial infection; autoimmune disease; diabetes;
 KW atopic dermatitis.
 XX Homo sapiens.
 OS Homo sapiens.

PN WO200274961-A1.
 XX 26-SEP-2002.
 XX 14-MAR-2002; 2002WO-US005109.
 XX PF 15-MAR-2001; 2001US-00810173.
 XX (HYSEQ-) HYSEQ INC.
 XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
 XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 XX Wehrman T, Wang J, Wang D, Dmanac RT;
 XX WPI; 2003-040556/03.
 DR N-PSDB; ABX05095.
 PT New isolated polypeptides and polynucleotides, useful for preventing,
 PT treating or ameliorating medical conditions, such as cancer,
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
 PT disorders, and infections.
 PS Claim 9; SEQ ID NO 636; 235pp; English.
 XX The invention relates to human polynucleotides and the polypeptides they
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,
 CC forensics, gene mapping, medical imaging, identification of mutations,
 CC responsible for genetic disorders or other traits, assessing biodiversity
 CC and producing many other types of data and products dependent on DNA and
 CC amino acid sequences. They are also useful for preventing, treating or
 CC ameliorating medical conditions, such as cancer, neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
 CC Sequences ABG9988-ABG9989 and ABU0010-ABU0043 represent human
 CC polypeptides of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied by the European Patent Office
 XX
 XX Sequence 418 AA;
 SQ
 Query Match 83.7%; Score 1917; DB 6; Length 418;
 Best Local Similarity 81.4%; Pred. No. 7.4e-18;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
 QY 1 MFSPGQEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKSRFPALYKRPANGVKSTVH 60
 DB 1 MFSPDQENH--PSKAPVKYIGELVVLGYNGSLPNGDRGRKSRFPALYKRPANGVKSTVH 58
 QY 61 VISTPQASKAISCKGSHISYTLSSRNOTVVEYTHDKDTMFPVGRSTESPIDFVYDTI 120
 DB 59 IACTPQAKAISNDQHSISYTLSSRAQTVEYTHDSNTDMFQIGRSTESPIDFVYDTV 118
 QY 121 SSGSNDDEAQTGSTIRFACRIVCDNREPTARIPAGPDSKNITLGEKAAKMKPDG 180
 DB 119 PGSQNSDTSVOSTISRFACRIICERNPPETARIVAAAGPDSKNITLGEKAAKMKTSDG 178
 QY 181 HMDGLTNGVLMVMPRGSGFTESQPGVWRISVGVGYTLRETSAAORGVLVESNTVL 240
 DB 179 QMGGLTNGVLMVMPRGFTEDSKFGIMREISVCGANFSLRETRSAQRGMVEIETNOL 238
 QY 241 QDSGLIDLCATLLMTADGLFHTPTOKHIEALRQEIINARPOCPVGLANTLAPESINRKE 300
 DB 239 QDSGLIDLCATLLMTABGLSHTPTVKHLEALRQEIINARPOCPVGFNTLAFPSMRKD 298
 QY 301 VVEKQWATLSCGHVGHYNNWGRSDTEANERCCPCRYGPPVPLMLGCEAGFYVDAG 360
 DB 299 VVDEKQWVYVINCCHVGHYNNWKEBERDGRCEPCRSYGPVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGAVGSEKAKYWSQIPLPHGTAFHFAACPPCATQLVGBONCKILFQGPID 420
 DB 359 PPTHAFPCGAVGSEKAKYWSQIPLPHGTHTFHAACPPCAHQVLAGEQGITRLIFQGPID 418

RESULT 13
 ADF83103
 ID ADF83103 standard; protein; 418 AA.
 XX
 XX ADF83103;
 AC
 XX 26-FEB-2004 (first entry)
 DT
 XX
 XX Human Pellino homologue 1, overexpressed in cancer.
 DE
 XX Human; Pellino homologue 1; Pellinol; cancer; cytostatic; vaccine;
 KW gene therapy.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO2003100000-A2.
 PN
 XX 04-DEC-2003.
 PD
 XX
 XX 22-MAY-2003; 2003WO-US016049.
 XX
 XX 24-MAY-2002; 2002US-0382606P.
 PR 25-JUL-2002; 2002US-0398099P.
 XX
 XX (TULA-) TULARIK INC.
 XX
 XX Li J, Ma D, Yang J;
 PI
 XX WPI; 2004-035118/03.
 DR N-PSDB; ADF83102.
 DR GENBANK; NP_065702.
 XX
 XX Diagnosing a cancer in a mammal comprises determining RecQ5, CTXL,
 PT USP13, MCL1, or Pellino 1 gene copy number in a biological sample from a
 PT region of the mammal that is suspected to be precancerous or cancerous..
 XX
 XX Claim 179; SEQ ID NO 10; 174pp; English.
 PS
 XX The present sequence is the protein sequence of human Pellino homologue 1
 CC (Pellinol), a protein associated with the kinase domain of activated
 CC Pelle. The invention is based on the finding of the overexpression of
 CC Pellinol and other genes (RecQ5, CTXL, USP13 and MCL1) in certain
 CC cancers, including breast cancer, colon cancer, lung cancer and ovarian
 CC cancer, and the frequent amplification of these genes in cancer cells..
 CC The genes, and their expression products, can be used diagnostically or
 CC as targets for cancer therapy. They can also be used to identify and
 CC design compounds useful in the diagnosis, prevention and therapy of
 CC tumours and cancers, in vaccine development, and in methods for
 CC determining the efficacy of a treatment regime. A claimed method for
 CC inhibiting cancer or precancerous growth, especially in ovarian or lung
 CC tissue, uses an inhibitor that interacts with Pellinol DNA or RNA. The
 CC inhibitor is a small interfering RNA (siRNA), microRNA (miRNA), an
 CC antisense RNA, and antisense DNA, a decoy molecule, a decoy DNA, a
 CC ribozyme or small molecule.
 XX
 XX Sequence 418 AA;
 SQ
 Query Match 83.7%; Score 1917; DB 8; Length 418;
 Best Local Similarity 81.4%; Pred. No. 7.4e-18;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
 QY 1 MFSPGQEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKSRFPALYKRPANGVKSTVH 60
 DB 1 MFSPDQENH--PSKAPVKYIGELVVLGYNGSLPNGDRGRKSRFPALYKRPANGVKSTVH 58
 QY 61 VISTPQASKAISCKGSHISYTLSSRNOTVVEYTHDKDTMFPVGRSTESPIDFVYDTI 120
 DB 59 IACTPQAKAISNDQHSISYTLSSRAQTVEYTHDSNTDMFQIGRSTESPIDFVYDTV 118
 QY 121 SSGSNDDEAQTGSTIRFACRIVCDNREPTARIPAGPDSKNITLGEKAAKMKPDG 180

DB 119 PGSQNSDTSVOSTISRFACRIICERNPPTARIYAAGFSDSKNIFLGEKAKMKTSDG 178
 QY 181 HMDGLTTNGVLMHPRGFTFEESOPGVWREISVCGDYTLRETRSAOQRKLVESNTNV 240
 DB 179 QMGGLTTNGVLMHPRGFTFEESOPGVWREISVCGDYTLRETRSAOQRKLVESNTNV 238
 QY 241 QDGSLLDLCGATLLMTADGLFHTPTQKHEALROEINARPOCPVGLTAPPSINRKE 300
 DB 239 QDGSLLDLCGATLLMTADGLFHTPTQKHEALROEINARPOCPVGLTAPPSINRKE 298
 QY 301 VVEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
 DB 299 VDEKQPMVYVLCNGHVGYNMGNKBERDCKRCEPMCRSVGPVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFTPCGHCVCSEKSAKYSQIPLPHGTHAFHAACPCATQVGEONCIKILFQSPID 420
 DB 359 PPTHAFTPCGHCVCSEKSAKYSQIPLPHGTHAFHAACPCATQVGEONCIKILFQSPID 418
 RESULT 14
 ID ADO44005 standard; protein; 418 AA.
 AC ADO44005;
 DT 15-JUL-2004 (first entry)
 XX Amino acid sequence of human Pellino 1.
 DE
 XX protein complex; neurological disease; stroke; neurodegeneration;
 KM Wallerian degeneration; Alzheimer's disease; neurological disorder;
 KM epilepsy; inflammatory condition; ulcerative colitis; Crohn's disease;
 KM atherosclerosis; ID-MYO-inositol triphosphate 3 kinase A; ASK1; ASK2;
 KM ASK3; Cankit beta; Cankit delta; Cankit gamma; casein kinase II alpha;
 KM Cdc37; CHK2; CTCL tumour antigen SE20-4; EF-1 alpha 1; EMAP;
 KM FLJ14653 NTRK2P2002252; FLJ30839 FEBRA2002429; HERC2;
 KM inositol polyphosphate-5-phosphatase; inositol-1; 4;
 KM 5-triphosphate 5-phosphatase type 1; IRAK1; IRAK4; KIAA1441; MSTP030;
 KM Nek9; PA3; Pellino 1; Pellino 3; podocalyxin-like protein 1 precursor;
 KM Pushover; S-adenosylhomocysteine;
 KM secretory carrier-associated membrane protein 2; surfactant protein 2;
 KM ubiquitin carboxyl terminal hydrolase 11;
 KM upstream regulatory element binding protein 1; Vartul;
 KM Werner's syndrome helicase interacting protein; WHIP;
 KM X-ray repair cross complementing protein 4.
 KM
 OS Homo sapiens.
 PN WO2004031242-A2.
 PD 15-APR-2004.
 PF 11-SEP-2003; 2003WO-EP010110.
 PR 12-SEP-2002; 2002EP-00020495.
 PR 12-SEP-2002; 2002EP-00020496.
 PR 12-SEP-2002; 2002EP-00020497.
 PA (CELL-) CELLZOME AG.
 PI Boummeester T, Drewes G, Jackson D, Helftenbein G, Schirle M;
 PI Kuester B, Hopf C;
 DR WPI; 2004-316467/29.
 XX
 XX New complex comprising at least one first protein, and at least one
 PT second protein, useful for treating stroke, Alzheimer's disease,
 PT neurological disorders such as epilepsy, and inflammatory conditions such
 PT as ulcerative colitis.
 PS Example; Page 224-226; 287pp; English.
 XX The specification describes protein complexes involved in cellular

CC processes which have been shown to be critical for the development of
 CC various forms of neurological diseases. Three protein complexes were
 CC identified: ASK2 protein complex, Pellino-1 protein complex and Pellino-3
 CC protein complex. The protein complex are useful for treating diseases and
 CC disorders, e.g. stroke, neurodegeneration such as Wallerian degeneration,
 CC Alzheimer's disease, neurological disorders such as epilepsy, and
 CC inflammatory conditions such as ulcerative colitis, Crohn's disease or
 CC atherosclerosis. Proteins identified as being part of the protein
 CC complexes of the invention are ID-MYO-inositol triphosphate 3 kinase A,
 CC ASK1, ASK2, ASK3, Cankit beta, Cankit delta, Cankit gamma, casein kinase
 CC II alpha, Cdc37, CHK2, CTCL tumour antigen SE20-4, EF-1 alpha 1, EMAP,
 CC FLJ14653 NTRK2P2002252, FLJ30839 FEBRA2002429, HERC2, two hypothetical
 CC proteins of 35.5 kDa and 49.3 kDa, inositol polyphosphate-5-phosphatase,
 CC inositol-1,4,5-triphosphate 5-phosphatase type 1, IRAK1, IRAK4, KIAA1441,
 CC MSTP030, Nek9, PA3, Pellino 1, Pellino 3, podocalyxin-like protein 1
 CC precursor, Pushover, a putative S-adenosylhomocysteine, secretory
 CC carrier-associated membrane protein 2, surfactant protein 2, ubiquitin
 CC carboxyl terminal hydrolase 11, upstream regulatory element binding
 CC protein 1, Vartul, Werner's syndrome helicase interacting protein (WHIP),
 CC X-ray repair cross complementing protein 4 (isoform 1). The present
 CC sequence represents Pellino 1.
 CC
 SQ Sequence 418 AA;
 Query Match 83.7%; Score 1917; DB 8; Length 418;
 Best Local Similarity 81.4%; Pred. No. 7, 4e-188;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
 QY 1 MESPQGEHCAPKKEPVYKIGELVYLGNGALPNQDGRKRSRALKRPKANGVKSSTN 60
 DB 1 MESPQGEHCAPKKEPVYKIGELVYLGNGALPNQDGRKRSRALKRPKANGVKSSTN 58
 QY 61 VISTPQASAKISCKGSHSISTLSRNQTVVEETHDKDDMPVGRSTSPIDFVYTDIT 120
 DB 59 IACTPQAKAISCKDHSISTLSRNQTVVEETHDKDDMPVGRSTSPIDFVYTDIT 118
 QY 121 SSGSONTDEAQITOSTISRFACRIICERNPPTARIYAAGFSDSKNIFLGEKAKMKTSDG 180
 DB 119 PGSQNSDTSVOSTISRFACRIICERNPPTARIYAAGFSDSKNIFLGEKAKMKTSDG 178
 QY 181 HMDGLTTNGVLMHPRGFTFEESOPGVWREISVCGDYTLRETRSAOQRKLVESNTNV 240
 DB 179 QMGGLTTNGVLMHPRGFTFEESOPGVWREISVCGDYTLRETRSAOQRKLVESNTNV 238
 QY 241 QDGSLLDLCGATLLMTADGLFHTPTQKHEALROEINARPOCPVGLTAPPSINRKE 300
 DB 239 QDGSLLDLCGATLLMTADGLFHTPTQKHEALROEINARPOCPVGLTAPPSINRKE 298
 QY 301 VVEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
 DB 299 VDEKQPMVYVLCNGHVGYNMGNKBERDCKRCEPMCRSVGPVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFTPCGHCVCSEKSAKYSQIPLPHGTHAFHAACPCATQVGEONCIKILFQSPID 420
 DB 359 PPTHAFTPCGHCVCSEKSAKYSQIPLPHGTHAFHAACPCATQVGEONCIKILFQSPID 418
 RESULT 15
 ID ADP48667 standard; protein; 418 AA.
 AC ADP48667;
 DT 09-SEP-2004 (first entry)
 XX
 XX Human Pellino-1 protein SEQ ID NO:4.
 DE
 XX nuclear factor kappa B dependent transcription inhibitor;
 KM NF-kB-dependent transcription inhibitor; Pellino-1;
 KM p38-dependent transcription inhibitor; Pellino-1;
 KM interleukin 1 receptor-associated kinase 4;
 KM IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
 KM antiarthritic; antiasthmatic; antiinflammatory; antirheumatic; antiulcer;

This Page Blank (uspto)